DOCTOR OF PHILOSOPHY

Spatio-temporal Modelling of Gene Regulatory Networks

Mitchell, Elaine Irwin

Award date:
2018

Spatio-temporal Modelling of Gene Regulatory Networks

By

Elaine Irwin Mitchell

A thesis submitted for the degree of Doctor of Philosophy

Division of Mathematics

School of Life Sciences

University of Dundee

Dundee

January 2018

Contents

Declaration

I declare that the following thesis is my own composition and that it has not been submitted before in application for a higher degree.

Elaine Mitchell

Certification

This is to certify that Elaine Mitchell has complied with all the requirements for the submission of this Doctor of Philosophy thesis to the University of Dundee.

Prof. Mark A. J. Chaplain

Prof. Sonia Rocha

Chapter 1

Introduction

Systems in biology, including ecosystems, populations, multicellular organisms, down to subcellular systems are comprised of many different networks on many different spatial and temporal scales. For example, our bodies have a variety of biological networks, for instance, the neural networks of the central nervous system; the cardiovascular system; the lymphatic system and on a sub cellular level, there are networks of gene regulatory networks (GRNs). The dynamics of these networks give rise to clocks that follow rhythms or oscillations with periods ranging from seconds to days. Our hearts beat every second, and cells divide periodically. The circadian clock located in the hypothalamus generates twenty-four hour rhythms, timing our sleep and hormone release. These time keeping strategies are underpinned at the sub cellular level by the oscillatory behaviour of GRNs (Bell-Pedersen et al. 2005).

Experimentalists attempt to unveil the complex mechanism behind these systems. However, experiments, can only probe so far depending on the technological capabilities. Additionally, the high complexity of these systems and their coupled non linear interactions, overthrows the power of human intuition. There has been great progress in mathematics to model such systems.

Sir D'Arcy Thompson was one of the pioneers of mathematical biology. His investigations of pattern formation in nature, (Thompson 1917), led to the introduction of appropriate mathematical models for reaction-diffusion systems (Turing 1952), with significant implications in genetics and development (Morriss-Kay & Sakolova 1996).

In addition, methods of stochastic processes have been developed and applied to a variety of problems in all areas of biology, but particularly in both population genetics and molecular genetics (Feller 1950, Fisher 1958, Lander et al. 2004, Skorokhod et al. 2002). There are many more sophisticated mathematical results that have contributed to, and benefited from, investigations in biology.

Experimentalists and mathematicians have began to collaborate and embrace the benefits of mathematical modelling and its predictive power. Mathematicians have began to model these complex systems and seek to probe solutions to help predict further functions of the system, (facilitating the use of computational power) and therefore, guide the experimentalists next hypothesis. Qualitative and quantitative approaches have been taken. Describing systems in a quantitative manner means their behaviour can be better simulated, and hence properties can be predicted that might not be evident to the experimenter. As computational methods developed and computational memory increased, the ability to solve complex non-linear systems has become possible. There has also been increasing interest in in-silico experimentation due to ethical considerations, risk, lack of material for multiple replication, unreliability and other complications involved in human and animal research.

Since the human genome project, an explosion in data-rich information sets have been uncovered. The human genome contains 20-25 thousand genes. Without the use of the analytics tools of mathematics, genomics information would be excruciatingly painful to understand. Additionally, systems biologists have been working to try and delineate the patterns of interactions between genes via mathematical modelling of the cell's

gene regulatory networks. This thesis will focus in on the intracellular level, mathematically modelling GRNs, seeking insight to their spatial-temporal dynamics.

Our genes and the networks of interactions between them, provide our cells with information on how to function and survive. Synthetic biologists and systems biologists work collaboratively in this highly interdisciplinary field to unfold the greatly yet to be understood biomolecular logistics and interactive communication mechanisms behind them. Together they have created and modelled synthetic GRNs in order to try and understand GRNs (Elowitz & Leibler 2000, Gardner et al. 2000).

Just one single gene can be known to interact with hundreds of other genes. However, there are sub networks that seem to appear commonly within the complexity of the overall cell genome. These subnetworks are called motifs Milo et al. (2002), Shen-Orr et al. (2002). We define network motif as patterns of interconnections that recur in many different parts of a network at frequencies much higher than those found in randomized networks. Positive and negative feedback loops make up two types of motifs. These motifs produce both spatial and temporal oscillations in molecular number.

Positive feedback loops generally aid in amplifying a signal, however, when combined together with a negative feedback loop, robust oscillations have been seen to be produced Nelson et al. (2004), Sneppen et al. (2010), Milo et al. (2002), Shen-Orr et al. (2002), Yeger-Lotem et al. (2004), Alon (2007).

The presence of the negative feedback loop, whether in the form of direct transcriptional repression or transcription factor sequestration is the ultimate key to sustaining oscillations Kim (2015). However, when combined with positive feedback, more robust oscillations have been seen to be produced Mitari et al. (2015), Tian et al. (2009).

GRNs which contain negative feedback loops are known to exhibit spatio-temporal oscillatory dynamics of the components involved Dolmetsch et al. (1997), Dolmetsch

et al. (1998), Berridge et al. (2003), Dequeant et al. (2006), Shimojo et al. (2008), Hirata et al. (2002), Nelson et al. (2004), Lahav et al. (2004).

Intracellular, spatial-temporal mathematical models can now be developed, validated and has been advancements in technology that can now visually observe the intracellular spatial-temporal dynamics of molecules, such as fluorescent imaging.

A negative feedback loop occurs when the output of a gene inhibits or represses the output of another gene in that network. Similarly, a positive feedback loop, is where the products of a gene enhances the products of another gene in the same network.

The most simple example being where the output product of a gene directly or indirectly up/down regulates its own output product.

The next step in cellular fate depends on the the cells response to the current state of its own genetic networks and microenvironment, and any stimuli from the surrounding environment.

The successful transfer and processing of all of this information can be greatly exhibited by the spatial and temporal dynamics of the cells molecular profile, which together then helps to trigger the next step.

More evidence is being acquired which points toward the dynamics of GRNs playing the major role in the next step of cell fate decision making. For example the p53/Mdm2 GRN is activated under DNA damage, exhibiting oscillations which aid to try and protect the cell (Lahav et al. 2004, Purvis et al. 2012); Hypoxia Inducible Factor-α (HIF-α)/Vascular Endothelial Growth Factor (VEGF) GRN is activated to regulate angiogenesis (Forsythe et al. 1996, Hirota & Semenza 2006); Notch signalling determines proliferation or differentiation in neurons Shimojo et al. (2008); Nuclear Factor κB (NFκB) is activated in the proinflammatory and immune responses (Ashall et al. 2009, Lee et al. 2014, Nelson et al. 2004); The Hes1 GRN helps to determine the outcome

of embryonic stem cell differentiation (Kobayashi & Kageyama 2011, Kobayashi et al. 2009); the HIF-α/prolyl hydroxylase (PHD) inhibits cell death under hypoxic conditions Bagnall et al. (2013).

GRNs which contain negative feedback loops (circuits) are known to exhibit spatio-temporal oscillatory dynamics of the components involved (proteins, mRNA). These oscillations may play a major role in crosstalk with other pathways (Veening et al. 2008, Wang et al. 2011) or assist steps in epigenetic modifications Nelson et al. (2004) and hence, co-activator/repressor recruitment and chromatin remodelling: each in part determining future gene activity.

In consideration of spatial pattern formation in embryonic development, it is quite intuitive that spatial models are essential.

Also, there has been experimental evidence showing that the spatial distribution of transcription factors within Eschericia coli is greatly determined by the position of the genes Kuhlman & Cox (2012).

At the intracellular level, we explore the spatio-temporal dynamics of different gene regulatory networks (GRNs), particularly different spatial dynamics at early and later times and how these early and late dynamics change as a result of a change in spatial parameters, such as diffusion coefficient and location of gene sites within a cell.

Chapter 2

Biological Background

In this chapter we review the biology of intracellular signal transduction. We focus on the spatial structure of eukaryotic cells, intracellular molecular transport, gene regulatory networks, negative feedback loops and stochasticity.

2.1 The spatial structure of the Eukaryotic cell

All multicellular organisms are made up of eukaryotic cells, though some eukaryotes exist as single cell organisms. A eukaryotic cell is contained within a plasma membrane. Eukaryotic cells contain an internal cytoplasmic compartment and nucleus within the cytoplasm. Cytosol, the jelly-like substance within the cell, provides the fluid medium necessary for biochemical reactions. Eukaryotic cells also contain various internal compartments called organelles, the nucleus being the central organelle of the cell. An organelle (”little organ”) is one of several different types of membrane-enclosed bodies in the cell, each performing a unique function. Organelle function can be divided into three main categories: information processing, energy processing, and packaging of chemical products. The nucleus is the organelle most associated with

information processing. Nuclear pore complexes determine where species can move between the nucleus and cytoplasm and also the rate at which this can happen. The nucleus contains long, single strands of DNA called chromosomes.

2.1.1 The cytoskeleton

The shape of a eukaryotic cell is maintained by the cytoskeleton. The cytoskeleton is a group of fibrous proteins that uphold cellular architecture, are responsible for organelle positioning and resist compression of the cell. However, they are flexible and can reshape the cell and relocate cellular organelles as required. The cytoskeleton is also important for cell reproduction, mechanotransduction and transportation of substances within the cell (Wang et al. 1993, Wirtz & Dobbs 1990). Protruding out from an organelle, known as the microtubule organising centre (MTOC) are many roads of filaments, called microtubules. The vehicles that transport many different molecules along this road network are called motor proteins: myosin, dynein and kinesin.

2.1.2 Molecular Transport

All cells are highly organised spatially. To help establish and maintain uneven distributions of specific molecules and organelles, eukaryotic cells employ several distinct mechanisms for molecular transport. The appropriate subcellular localisation of molecular species is critical for a cell to remain healthy (Johansson et al. 2008, Kim et al. 2000, Norvell et al. 2005).

Diffusion

Diffusion is the main mechanism of transport for many important materials in the cell. It is sometimes described as 'passive transport' as it does not require any energy. Simply put, it is the movement of molecules from a region of high concentration to a region of low concentration: molecules moves with their concentration gradient.

Active transport

Unlike diffusion, the active transport of material is directed and requires energy. The term active transport can describe molecular transport across a membrane or the transport facilitated by motor protein along the microtubules of the cytoskeleton.

Depending on the motor protein, trafficking of intracellular molecules can be directed from the nuclear membrane to the cell membrane or in opposite direction. Dynein motors attach to proteins and transport them along microtubules towards the nucleus and kinesin motors attach to proteins and transport them along microtubules away from the nuclear membrane. Active transport along microtubules is involved in information transfer as it helps translocate molecules to and away from the nucleus (quickly), where the information processing occurs.

2.2 Signal transduction

The ability to sense and adjust to the environment is crucial to life. For multicellular organisms, the ability to respond to external changes is essential not only for survival but also for normal development and physiology. Cells have a communication network, underpinned by genetic circuitry.

Intracellular signal transduction can be described as the transmission of extracellular signals from a cell's exterior to its interior, for example, to the nucleus. Molecular signals are transmitted between cells by the secretion of hormones and other chemical factors. The ability of an organism to function normally is dependent on all the cells of its different organs communicating effectively with their surrounding environment and with each other - a phenomenon known as intercellular communication. Mammalian cells require stimulation for cell division and survival; in the absence of growth factor, the cell will undergo apoptosis (cell death). These extracellular stimulation requirements are necessary for controlling cell behaviour in unicellular and multicellular organisms. In fact, signal transduction pathways are perceived to be so central to biological processes that a large number of diseases are attributed to their dysregulation.

Signal transduction pathways are very specific, both in their activation and in their downstream actions. Thus, the various organs in the body respond in an appropriate manner and only to relevant signals. In eukaryotic cells, most intracellular proteins are activated by a ligand/receptor interaction and possess an enzymatic activity, examples include tyrosine kinases and phosphatases. This can start off a cascade of reactions, activating further molecules to respond, each acting as a messenger in the process. These secondary messengers then bring into play complex gene regulatory networks (GRNs) which respond appropriately and decide the next step in cell decision making.

2.2.1 Gene expression

Segments of DNA (deoxyribonucleic acid), called genes, are the building block of life. Genes encode all the information our cells need to communicate, function, grow and survive. Depending on cell type, for instance, muscle or neuronal, a different set of genes (20 - 25 thousand for human beings) will be functional.

A gene can be activated/deactivated by specific signalling molecules (gene regulators) called transcription factors. Activation/deactivation occurs due to specific transcription factors binding to regulatory regions of the gene, namely promoters, enhancers and silencers. These regulatory regions, known as boxes, are specific repeating sequences of nucleotides. For example, an N box denotes a conserved promoter element and an E box denotes an enhancer element.

Once this transcription factor-regulatory region binding occurs, enzymes, (biological catalysts), called RNA polymerase are able/unable to bind to the targeted gene and initiate a process called transcription. Transcription is the process by which the information encoded in DNA is copied into a new molecule of messenger RNA (mRNA). Hence, the process being mediated by transcription factors, determining the factor at which an mRNA transcript can be produced. Some genes are located closer to the nuclear membrane than others, increasing their sensitivity to transcription factors (Cole & Scarcelli 2006).

mRNA provides a template to the synthesis of a protein. A process called alternative splicing of mRNA will result in a different protein being synthesised. Once an mRNA is formed, it can travel to the nuclear membrane, be actively transported across it and diffuse in the cytoplasm to reach ribosomes where its associated protein can be synthesised.

Proteins can be classified depending on their domains. The transcription factor called Hes1 contains the basic helix-loop-helix (bHLH) domain and NF-κB contains the Rel homology domain. These domains (along with other aspects, such as protein folding) specify the functional purpose of the protein. Particularly, it specifies, which genetic regulatory sites the protein can bind to.

Once formed, a transcription factor, with an exposed nuclear localisation sequence (NLS) can travel across the nuclear membrane into the nucleus to meet its associated

target genes. Once again, transcription factor-regulatory region binding can induce or inhibit (depending on the functional role) the transcription of mRNA encoded in that gene, thus, controlling levels of mRNA and protein copy numbers.

This is a description of gene interaction.

2.2.2 Gene Regulatory Networks

A gene regulatory network (GRN) lies at the core of intracellular signal transduction. In brief, a GRN is a collection of genes in a cell which interact with each other indirectly through their RNA and protein products (and with other substances in the cell), thereby governing the rates at which genes in the network are transcribed into mRNA.

If the products of a first gene enhance or inhibit the products of a second gene, we say that the first gene positively or negatively regulates the second gene. If the output of the second gene feeds back in to positively up regulate or negatively down regulate the first gene, we call this circuitry a positive/negative feedback loop.

If the output product of a gene directly or indirectly up/down regulates its own output product, we say that the GRN contains an auto activating/inhibiting feedback loop.

Patterns of protein-protein or transcription factor-gene interactions appear consistently throughout many cell types. It has been found that these patterns of interconnections tend to have a functional information processing purpose. These patterns are called regulatory motifs and are defined more specifically as patterns of interconnections that recur in complex networks significantly more often than in randomized networks. Positive and negative feedback loops are commonly found in regulatory motifs Sneppen et al. (2010), Milo et al. (2002), Shen-Orr et al. (2002), Yeger-Lotem et al. (2004), Alon (2007).

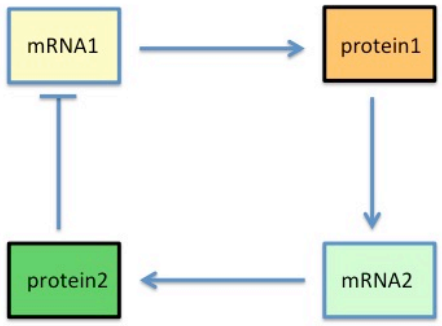

Figure 2.1: *A simple schematic example of a negative feedback loop:gene one activates gene 2 and in turn gene 2 inhibits gene 1*

More precisely, GRNs which contain negative feedback loops are known to exhibit spatio-temporal oscillatory dynamics of the components involved (proteins, mRNA) Dolmetsch et al. (1997), Dolmetsch et al. (1998), Berridge et al. (2003), Dequeant et al. (2006), Shimojo et al. (2008), Hirata et al. (2002), Nelson et al. (2004), Lahav et al. (2004).

It has been found that networks containing solely negative feedback can produce coherent oscillations, although, period of oscillations have been found to be very diverse.

Positive feedback loops generally aid in amplifying a signal, monostability, bistability. However, when combined together with a negative feedback loop, more robust oscillations have been seen to be produced than a network containing only negative feedback Mitari et al. (2015), Tian et al. (2009).

Figure 2.3 shows a schematic diagram of a GRN, where each node represents a gene

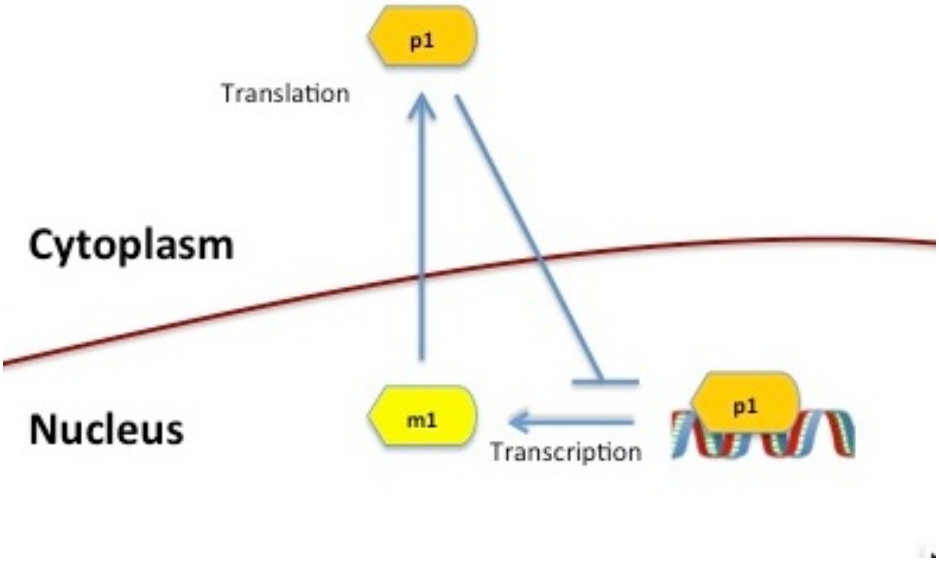

Figure 2.2: *A simple schematic of the Hes1 GRN: One gene repressilator*

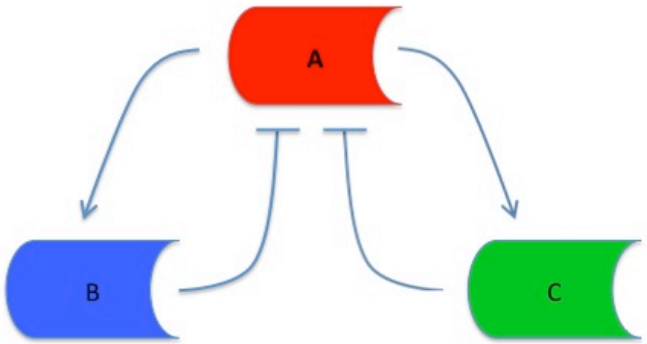

Figure 2.3: *Negative feedback GRN: The products of A enhances output of both B and C, which in turn downregulate output production from A*

of the GRN (i.e. gene A, B or C) and each arrow indicates the positively or negatively regulated pathway due to that genes output (mRNA and/or protein). Thus, the output products of gene A induces the output products of both B and the output products of C, to which the output products of both B and C in turn inhibit and hence negatively regulate the productive activity of A.

The digital information provided by multiple extracellular signals is rapidly decoded, processed and encoded as an analogue intracellular signal Kyriakis (2007), Jaenisch & Bird (2003), Mohan (2009), Behar & Hoffmann (2010), Behar & Hoffmann (2013), Purvis & Lahav (2013), Tay et al. (2010).

It is believed that analogue information, such as frequency, amplitude and response time of transcription factor spatial-temporal abundance have important implications in gene expression Tay et al. (2010), Paszek et al. (2010), Mengel et al. (2007).

It is believed that the oscillations seen, carry information that play a major role in determining cell fate Ashall et al. (2009), Purvis et al. (2012), Lahav et al. (2004). This may be due to the oscillations playing a major role in crosstalk with other pathways Wang et al. (2011), or assist steps in epigenetic modifications Nelson et al. (2004) and hence, co-activator/repressor recruitment and chromatin remodelling: each in part determining future gene activity.

A major step in predicting the fate and phenotype of a cell lies in understanding the spatial temporal dynamics produced by these GRNs.

2.3 Stochasticity

Even in a population of isogenic cells in a homogeneous environment, the oscillations observed between cells are asynchronous if even oscillations are seen at all in

a particular cell. This heterogeneity has been widely recorded since the development of fluorescent imaging. It is believed that intrinsic and extrinsic stochasticity in the biochemical and transport dynamics of a GRN play a major role in this asynchrony and that it may have a purpose, specifically in areas of gene expression, biological development, cell survival and evolution Eldar & Elowitz (2010).

The work of Carr & Whitmore (2005) is concerned with zebrafish tissue and cell line containing circadian clocks. If light is deprived, population assays show a dampening of oscillations. However, individual cell imaging showed that individual cells still oscillated but with widely different phases from other cells. A pulse of light helped to synchronise the cell line once again.

Genetically identical cells and organisms exhibit remarkable diversity even when they have identical histories of environmental exposure. for example, the differences in identical twins finger prints or the difference in coat patterns and personalities in a cloned cat Raser & O'Shea (2005).

Noise, or variation, in the process of gene expression is now believed to contribute to this phenotypic variability.

Heterogeneity in gene expression has a substantial influence on susceptibility to disease, disease prognosis, the efficiency of drug therapy, and development of mental health disorders such as drug dependence Khokhar et al. (2010), Madian et al. (2012), Thuong et al. (2008).

Stochasticity in gene expression is believed to aid in survival at an evolutionary level of selection. Waddington's theories of canalization and genetic assimilation propose that wasteful phenotypic variability in a population is suppressed when the population is well adapted to its environment. However, if environmental conditions shift, phenotypic noise becomes advantageous because a noisy population will produce some

members that are better adapted to the new environment.

Most of the information on cell signalling has been obtained from population-level studies using bulk assays, yet it is not clear if population data faithfully reflect how individual cells respond. For example, pulsed responses of p53 to radiation damage are evident only at the single-cell level Lahav et al. (2004). The results of Paszek et al Turner et al. (2010) suggest that a stochastic threshold, controlling functional all-or-nothing responses, is behind the source of variability shown between identical individual cells in NFκB signalling.

Genes are organised into regulatory circuits where the expression of one gene can influence the expression of another. A consequence of this organisation is that noise in the expression of one gene may propagate to affect noise in the expression of a downstream gene

When large numbers of molecules are present, chemical reactions may proceed in a predictable manner, thus making PDE models sufficient in describing the spatial-temporal dynamics. However, when only a low number of molecules of a specific type exist in a cell, stochastic effects can become prominent and the PDE formulation, that calculates mean values, no longer captures the true variable dynamics.

Chapter 3

Mathematical Modelling of Gene Regulatory Networks

3.1 Introduction

As discussed in chapter two, a gene regulatory network (GRN) is a collection of genes within a cell, that interact directly or indirectly via their mRNA and protein output products. Patterns of protein-protein/transcription factor-gene interactions appear consistently throughout many GRNs in many cell types. It has been found that these patterns of interconnections tend to have a functional information processing purpose. These patterns are called regulatory motifs.

The following gives a few reasons to why mathematical modelling of GRNs is essential. GRNs are large and complex. There is extensive genomic information to decipher. There is still the search to decipher gene/protein interaction and to understand the consequential implications of minor changes to a gene or a protein. Mathematical models are required in order to gain further understanding of a GRN's activity, function and

purpose. Intuitive thinking of such complex processes is extremely limited and can only take an experimentalist so far.

Mathematical models have been utilised to recapitulate empirical information from the transcriptional code of a single gene (Yuh et al. 2001), to deduce previous unknown functional interactions from complex experimental data, leading the researcher to new testable hypotheses (Jaeger et al. 2004) and to predict features underpinning regulation of the system that are hard to measure directly (Fakhouri et al. 2010).

Mathematical models can capture the spatial-temporal oscillatory dynamics that a lot of GRNs produce via the presence of negative feedback (Berridge et al. 2003, Dequeant et al. 2006, Dolmetsch et al. 1997, 1998, Hirata et al. 2002, Lahav et al. 2004, Nelson et al. 2004, Shimojo et al. 2008). A negative feedback loop in a GRN is the situation where the output of a gene inhibits or represses the output of another gene in that network. The most simple example being where the output product of a gene directly or indirectly up/down regulates its own output product. Similarly, a positive feedback loop is the situation where the products of a gene enhances the products of another gene in the same network. Positive feedback loops generally aid in amplifying a signal. However, when combined together with a negative feedback loop, robust oscillations have been seen to be produced (Alon 2007, Milo et al. 2002, Nelson et al. 2004, Shen-Orr et al. 2002, Sneppen et al. 2010, Yeger-Lotem et al. 2004).

The presence of the negative feedback loop, whether in the form of direct transcriptional repression or transcription factor sequestration is the ultimate key to sustaining oscillations (Kim 2015). However, when combined with positive feedback, more robust oscillations have been seen to be produced (Mitari et al. 2015, Tian et al. 2009).

Figure 3.2 shows a schematic diagram of a GRN, where each node represents a gene of the GRN (i.e. gene A, B or C) and each arrow indicates the positively or negatively regulated pathway due to that genes output (mRNA and/or protein). Thus, the output

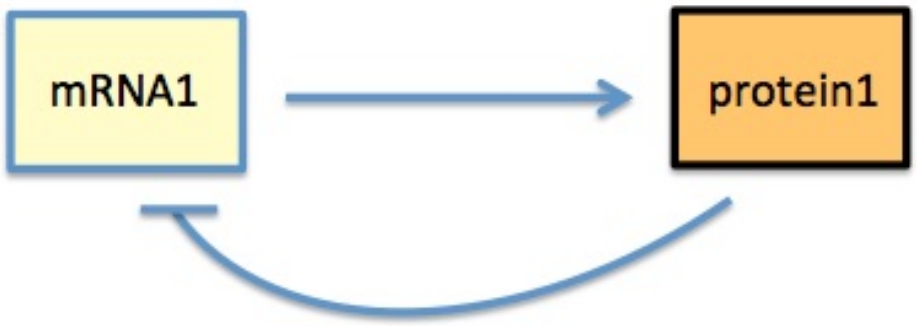

Figure 3.1: *A simple schematic diagram of the Hes1 GRN: The one gene repressilator.*

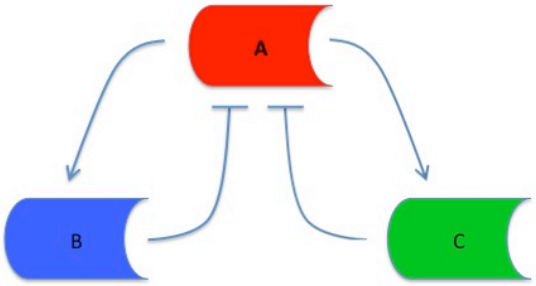

Figure 3.2: *A theoretical example of a negative feedback GRN: The products of gene A enhances output of both gene B and gene C, which in turn downregulate output production from gene A.*

products of gene A induces the output products of both B and the output products of C, to which the output products of both B and C in turn inhibit and hence negatively regulate the productive activity of A.

It is believed that the oscillations seen, carry information that play a major role in determining cell fate (Ashall et al. 2009, Lahav et al. 2004, Purvis et al. 2012). This may be due to the oscillations playing a major role in crosstalk with other pathways (Wang et al. 2011) or assist steps in epigenetic modifications (Nelson et al. 2004) and hence, co-activator/repressor recruitment and chromatin remodelling: each in part determining future gene activity.

A major step in predicting the fate and phenotype of a cell lies in understanding the spatial temporal dynamics produced by these GRNs.

3.2 A brief history of mathematical modelling of GRNs

Mathematical modelling of GRNs has been previously carried out and investigated extensively both from a deterministic and a stochastic modelling viewpoint. This overview will be broken up into separate sections focussing on deterministic and stochastic models respectively.

3.2.1 Deterministic models

One of the pioneers of mathematical modelling of GRNs was B.C. Goodwin who developed the first auto-inhibiting feedback loop, subsequently referred to as 'The Goodwin Oscillator' in the early 1960s and was seeking oscillatory dynamics of mRNA and protein concentration levels (Goodwin 1965). Similar work shortly followed (Griffith

1968), continuing to seek further insight into systems of closed, auto-inhibiting feedback loops. The GRNs modelled have the same network topology as the subsequently discovered actual Hes1 system (Hirata et al. 2002).

These models were deterministic, solely considering temporal dynamics and so were systems of ODEs. The following ODE is a model of the auto-inhibiting feedback loop.

$$\begin{aligned} \frac{dm}{dt} &= \frac{\alpha_m}{1+p^h} - \delta_m m, \\ \frac{dp}{dt} &= \alpha_p m - \delta_p p, \end{aligned} \tag{3.1}$$

where α_m describes the basal transcription rate, α_p describes the rate of translation at the ribosome, h is the hill function coefficient, describing the amount of cooperativity and δ_m, δ_p describe the rate of mRNA and protein degradation respectively

The oscillations observed were very sketchy. Such models do not take into account the spatial aspects of the cell, localised molecular synthesis and molecular movement, such as diffusion and active transport. Since biolochemical interactions are complex, the models used to describe GRNs contain high non linearities. Solutions without the use of the high level programming around today were completely limited, costly and inefficient. Thus, steady state solutions of ODEs were sought after.

A delay to account for transcription and translation was introduced (Mackey & Glass 1977). Shortly after Smolen, Baxter and Byrne embraced this addition (Smolen et al. 2001, 2002, 1999). In particular, Smolen et al. (1999) discussed the use of the delay to account for macromolecular transport. Specifically, their GRN model incorporated a delay to describe active transport of molecules and demonstrated that even though such a model resulted in oscillatory behaviour, the inclusion of molecular diffusion suppressed oscillations. Similar style GRN models have since been applied to the Wnt

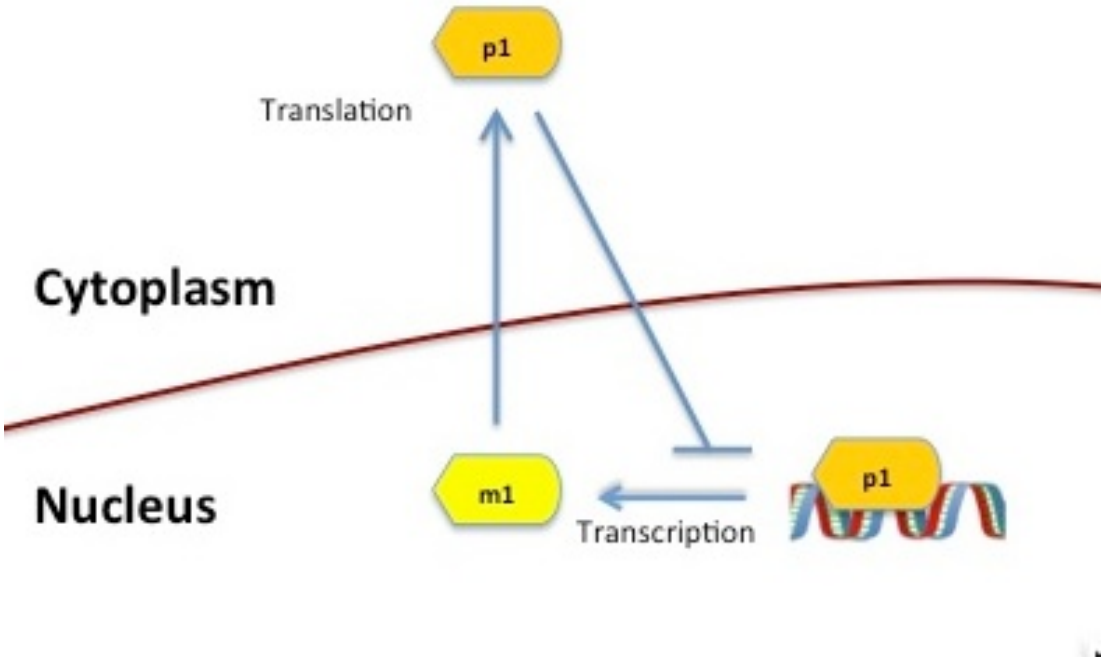

Figure 3.3: *A simple schematic of the Hes1 GRN: auto-inhibiting, negative feedback loop. Hes1 mRNA is transcribed at the gene site within the nucleus, it then translocates to the cytoplasm, where the ribosomes are located and protein synthesis can begin. Once Hes1 protein is synthesised, Hes1 protein can travel back to the nucleus and bind to its own Hes1 promoter and inhibit hes1 mRNA transcription from taking place, thus, down regulating its own production.*

system, Hes1 system, the p53-Mdm2 system and the NF-κB system, (Bernard et al. 2006, Jensen et al. 2003, Lewis 2003, Monk 2003, Tiana et al. 2002, Wawra et al. 2007), each model showing that delays were found to provoke oscillatory behaviour.

Theoretical models of synthetic GRNs (e.g. the repressilator) have also been proposed and studied (O'Brien et al. 2012, Purcell et al. 2010), while interest in modelling bacterial operons by Mackey and co-workers (Hilbert et al. 2011, Mackey et al. 2015, Yildirim & Mackey 2003) has added additional insight.

Furthermore, ODE models of GRNs were extended by the incorporation of transformation of nuclear species to cytoplasmic species (and vice versa) to facilitate molecular

movement between the nucleus and the cytoplasm (Agrawal et al. 2009, Lipniacki et al. 2004).

However, up until the advancements of fluorescent imaging of individual cells, most cell signalling information had been obtained from population-level studies using bulk assays. This experimental technique calculates averages across the entire cell (or nucleus and cytoplasm if specific methods are implemented), thus encouraging modellers to widely adopt ODE approach to modelling of GRNs.

Furthermore, modelling using ODEs assumes that the interior of a cell is spatially homogeneous (cf. well-stirred reactor) such that spatial movement of molecules (diffusion and active transport) and localisation of gene sites and molecular reactions etc are ignored. However, the interior of a eukaryotic cell exhibits a number of different spatial compartments e.g. nucleus, cytoplasm, golgi complex, where each compartment permits different biochemical activity and is usually enclosed by a thin lipid membrane. For the cell to remain healthy it is critical that molecular species remain localised to the appropriate compartments. Even the different compartments themselves are not spatially homogeneous e.g. the nucleus contains the nucleolus and chromosomes, and the crucial process of transcription of mRNA can only take place at specific gene sites on the chromosomes. Kuhlman & Cox (2012) found that the spatial distribution of protein was dependent not only on the spatial location of the genes but also on the state of DNA compaction. In fact, *E. coli* could further control the spatial distribution of protein by modifying the state of its nucleoid.

Mathematicians in the 1970s realised the importance of spatial models (Glass & Kauffman 1974, Shymko & Glass 1974), and again in the 1980s by Mahaffy and co-workers (Busenberg & Mahaffy 1985, Mahaffy 1984, 1988), where the focus was on analysing

generic systems with one-dimensional models. ODE models were reconfigured to incorporate a spatial dimension using reaction-diffusion PDEs and steady states and stability were determined with particular attention paid to the geometry of the model. They coined the term "spatial switching" to indicate how the system geometry can lead to different dynamical behaviour. This approach has recently been extended by Naqib et al. (2012). More recently, PDE models of GRNs have been extended to focus on actual GRNs (e.g. Hes1, p53-Mdm2, NF-κB) and in two spatial dimensions, (Sturrock et al. 2011*b*, Terry & Chaplain 2011).

3.2.1.1 Deterministic models of the n-gene repressilator

The system of reaction diffusion PDEs below, will be used throughout this thesis to model the deterministic, generalised n-gene repressilator GRN:

For $1 \leq i \leq n$,

$$\begin{aligned} \frac{dm_i}{dt} &= D_{m_i}\nabla^2 m_i + \frac{\alpha_{m_i}}{1+p_{i-1}^h} - \delta_{m_i} m_i \\ \frac{dp_i}{dt} &= D_{p_i}\nabla^2 p_i + \alpha_{p_i} m_i - \delta_{p_i} p_i, \end{aligned} \tag{3.2}$$

and for $i = 1$, we define $i-1 = 0 = n$. Here α_m describes the basal transcription rate, α_p describes the rate of translation at the ribosome, h is the hill function coefficient, describing the amount of cooperativity and δ_m, δ_p describe the rate of mRNA and protein degradation respectively.

For the special case, $n = 1$, the gene product represses its own output, i.e.

$$\begin{aligned} \frac{dm_1}{dt} &= D_{m_1}\nabla^2 m_1 + \frac{\alpha_{m_1}}{1+p_1^h} - \delta_{m_1} m_1, \\ \frac{dp_1}{dt} &= D_{p_1}\nabla^2 p_1 + \alpha_{p_1} m_1 - \delta_{p_1} p_1. \end{aligned} \tag{3.3}$$

We will term this case the one gene repressilator. The one-gene repressilator exists in cellular biology as the Hes1 GRN. O'Brien et al. (2012) discusses the ODE model: equation 3.1. This ODE model does not exhibit sustained oscillations but rather exhibits a short lived transient of damped oscillations leading to steady state solutions. With the inclusion of intermediate species, (incorporating for a delay), sustained oscillations were achieved, as experimentally observed by Hirata et al. (2002). Other delay ODEs exhibit oscillations (Momiji & Monk 2008, Monk 2003). Mather et al. (2009) showed that tight repressor binding and strong enzymatic decay of the protein were essential to obtaining oscillations.

Sturrock et al. (2011*b*) accounted for spatial aspects and found that oscillations for the PDE model of Hes1: equations 3.3, were achieved for a range of diffusion coefficients. They found that there was an upper and lower limit on the diffusion coefficient to which sustained oscillations could occur. Furthermore, it has rigorously been shown, (Chaplain et al. 2014) that sustained oscillations are generated by a limit cycle, where the diffusion coefficients of the molecules (Hes1 protein, hes1 mRNA) is the bifurcation parameter that gives rise to the limit cycle. Thus, in this thesis we will use a spatial model which explicitly includes the spatial movement of protein and mRNA molecules within a cell (nucleus, genes and cytoplasm).

The first ever synthetic GRN implemented in a living bacterium by Elowitz & Leibler (2000) is the GRN known as the original repressilator, comprising of a ring of three repressive genes. We refer to it in this thesis as the three-gene repressilator. Their experimental data captured oscillatory dynamics in the protein expression levels.

ODE models of the three gene repressilator, (Bennet et al. 2007, Elowitz & Leibler 2000) indeed show its oscillatory nature. It was shown there is a bifurcation parameter, which gives rise to a limit cycle (Muller et al. 2006, O'Brien et al. 2012, Smith 1987). Damped oscillations or steady states were achieved otherwise. The findings of Smith

(1987) and Muller et al. (2006) were extended to the generalised n-gene repressilator. They showed that the n-gene repressilator ODE model contained a globally attracting stable limit cycle for odd n and therefore behaves as a stable oscillator following the Hopf bifurcation. Thus, under the right parameter conditions, repressilators for n=3,5 etc were predicted to exhibit sustained oscillations.

Bennet et al. (2007) went further to look at the transient in the dynamics of the three-gene repressilator before the long term equilibria is reached. Lohmann et al. (2017) and Strelkowa & Barahona (2010, 2011) asserts that characterising the long-term attractors of the system might not be enough, since unstable solutions can play a significant role. Since, the time scales that are relevant in intra-cellular dynamics are finite, the relevant dynamics are not dictated exclusively by the long-term attractors of the system but can be dominated by reachable long-lived transients with qualitatively different observable dynamics.

For the case, n even, it was proved that the n-gene repressilator ODE system contains globally attracting steady state solutions exhibiting multi-stability, (Muller et al. 2006, Smith 1987). The n-gene repressilator, therefore, providing a natural bridge between oscillatory and steady state solutions. Thus, the long term attractors of the, $n = 2, 4, 6$ etc repressilator systems are steady state solutions. There were three possible fixed points of the system. One unstable, where all protein species were equal,

$$p_i = p_{i+1} = p_m, \quad \forall i. \tag{3.4}$$

The other two being a dimerised stable state, where,

$$p_i = p_{i+2} = p_{up} > p_m \tag{3.5}$$

$$p_j = p_{j+2} = p_{down} < p_m \quad i \neq j. \tag{3.6}$$

For $n = 2$, this prediction was supported by the first genetic switch implemented in *E. coli* by Gardner et al. (2000), corresponding to the two-gene repressilator. However, they began with the state in equilibrium, (where one gene is dominantly expressed and the other repressed). They then expose the system to an inducer of the repressed gene, after which the state of the system switches. The previous repressed gene became dominantly expressed while the previous dominantly expressed gene became repressed.

Strelkowa & Barahona (2011) put their focus on the n even case. They propose that the transient dynamics before equilibria is reached can not be ignored because the transient is long-lived, relative to the time-scales of intra-cellular dynamics. They show that the n-gene repressilator contains quasi-stable, oscillatory, finite-time transients that can not be ignored in such a transient and noisy environment as the cell (Lohmann et al. 2017). Their ODE models predict that the transient time of oscillations and the period in the n-gene (even and odd) repressilator increase with n.

Strelkowa & Barahona (2010) found that even rings have a quasi-stable limit cycle which, although unstable in terms of linear Floquet stability analysis, (stability in terms of analysis of the Poincare map), has only one unstable direction with a very slow escape rate. This means that trajectories are attracted to the limit cycle from all directions but one, hence leading to long-lived, inducible periodic transients in the deterministic setting and to sustained oscillations in the stochastic system However, they state that quasi-stability in this sense is a local property. They assert that to assess if these solutions will be reachable (and therefore observable in the dynamics), one needs to employ global techniques, e.g. sampling the space of initial conditions.

Stress response signalling pathways, such as the p53 and NF-κB pathways, have been found to respond to stimuli with transient oscillations (Kemler & Fontana 1999, Monk 2003). While the oscillations do not persist, the dynamical properties of the fluctuations can determine the specific downstream response (Hoffmann et al. 2002).

3.2.2 Stochastic models

Due to biochemical processes being intrinsically stochastic and the fact that the intracellular copy number of species, (such as mRNA and protein for example) is low, the continuum formulation of both ODEs, DDEs and PDEs does not fully capture the true dynamics experimentally observed. In a population of isogenic cells in a homogeneous environment, the oscillations observed are asynchronous, with a fraction of cells showing no activation. This heterogeneity has been widely recorded since the development of fluorescent imaging and the advancement in micro-assays. It is believed that intrinsic and extrinsic stochasticity in the biochemical and transport dynamics of a GRN play a major role in this asynchrony and that it may have a purpose, specifically in areas of gene expression, biological development, cell survival and evolution (Eldar & Elowitz 2010).

Stochastic simulations have therefore been used to address this randomness. These types of models are usually either systems of stochastic differential equations (Kepler & Elston 2001, Meister et al. 2014, Tian & Burrage 2006) or discrete stochastic simulations, such as Gillespie's algorithm (Gillespie 1977). Time delay versions of these methods have also been investigated (Bratsum et al. 2005, Ribeiro et al. 2006, Tian et al. 2007).

Moreover, the presence of noise in nonlinear systems may induce non-stationary dynamics in systems with only fixed point attractors in the deterministic setting (Suel et al. 2006) or, conversely, noise may act as a stabiliser of unstable deterministic states (Turcotte et al. 2008).

Meister et al. (2014) asserts that in multistable stochastic systems, large fluctuations can cause individual trajectories to escape from the domain of attraction of one steady

state solutions and be attracted to another, so the system eventually reaches a multi-modal probability distribution. Thus, multistable systems, although exhibiting static steady states in their long term dynamics can operate as switches in the stochastic regime.

Further to our argument toward using spatial models Becskei et al. (2005) showed that the position of genes on a chromosome in the nucleus and the affects of genetic activation events propagate further noise in the system, bearing consequences to GRN dynamics.

One of the first GRNs to be modelled with a spatial stochastic approach was the Min System in an Escherichia coli cell (Fange & Elf 2006, Howard & Rutenberg 2003). Howard & Rutenberg (2003) used a stochastic analogue of a one-dimensional system of reaction-diffusion equations and found that for some parameter values the protein concentrations were low enough that fluctuations were essential for the generation of patterns. In the model of Fange & Elf (2006), trajectories were generated using the next subvolume method (NSM), and numerical simulations were able to reproduce all documented Min phenotypes, where deterministic or non-spatial models could not. A spatial stochastic model of the MAPK pathway was developed in the study of Takahashi et al. (2010). This model was implemented numerically using a Green's function reaction dynamics scheme, which allows for individual particle-level simulation of molecular species. Using this technique, MAPK responses that could not be observed using a mean-field approach were simulated. Another recent spatial stochastic model was developed to study in detail a generic transcription factor binding and unbinding to DNA (van Zon et al. 2006). Here, the spatial stochastic model was able to support the use of well-stirred, zero-dimensional models for describing noise in gene expression. It is clear from these few examples that spatial stochastic modelling can provide insight into intracellular signalling pathways that other approaches cannot.

3.2.2.1 Stochastic models of the n-gene repressilator

Studies have investigated how time-delay in gene expression can cause a system to be oscillatory even when its deterministic counterpart exhibits no oscillations. Barrio et al. (2006) and Bratsum et al. (2005) show how oscillations can arise for the one-gene repressilator, due to the coupling of time-delay and noise. The stochastic simulation algorithm (SSA) of Gillespie (Gillespie 1977), was adapted to account for delay. These models found that oscillations persisted for a wider range of paramaters than their deterministic counterparts.

Recently, more and more experimental evidence in the form of bimodal population distributions indicates that noise plays a very important role in the switching of bistable systems, such as the two-gene repressilator (Tian & Burrage 2006). Deterministic models are not able to account for these fluctuation-induced transitions. Warren & ten Wolde (2005) found using mean field analysis and stochastic simulation of chemical models that the genetic toggle switch: two-gene repressilator, is more unstable (prone to switching) under the noise induced by gene expression. Wang et al. (2007) explored the switching affect with the addition of noise into their ODE model, specifically in the degradation rate of protein as did Xu et al. (2014). However, it is noted that this stochastic-switching affect is rarely seen in experiment, owing to the toggle switch being a robust system and this was confirmed by most stochastic models.

The pioneering work of Elowitz & Leibler (2000) configured the three-gene repressilator in *E. coli*. They introduced a temporal stochastic model to capture the dynamics and in agreement with deterministic models, found oscillations in the three protein species. The stochastic nature of the system captured a slight variability in period but also a rather large variation in protein level amplitude. The three protein species appear to oscillate with similar periods, yet out of phase with one another.

Potapov et al. (2015) lifted the assumption made of fast transcription factor-promoter binding and noted bimodal distributions for the period of oscillations of the three-gene repressilator. Analysis of the stochastic model used by Potapov et al. (2015) found a region in parameter space where both steady state and oscillatory dynamics could be found. They propose that due to the existence of this co-existing region (of which they call hysteresis), stochasticity could theoretically act as a switch between the two states.

Strelkowa (2011) does parallel analysis of deterministic and stochastic models of the n-gene repressilator. For n odd, she finds sustained oscillations in the dynamics. For n even, internal noise did not seem to destroy the quasi-stable oscillations seen in their deterministic model counterpart but that in fact, internal noise effectively prolonged the transient behaviour.

3.3 Spatial Stochastic Modelling of GRNs: General Approach

The development of mathematical models which reflect both spatiotemporal and stochastic aspects of GRNs can be regarded as an important computational tool in making predictions about the behaviours of GRNs and in the optimising of targeted drug treatment. In this thesis the spatial stochastic modelling approach that will be adopted is that used in Sturrock et al. (2013). We now describe this approach.

Cellular size is of the order of tens of micrometers. Thus, we are dealing with a mesoscopic system. For fine discretisations of the cellular domain, the classical SSA becomes inefficient. The NSM (Elf & Ehrenberg 2004) is an algorithm adapted for simulations of the reaction diffusion master equation (RDME), and it inherits good scaling properties from the Next Reaction Method (NRM) (Gibson & Bruck 2000).

In order to begin constructing our spatial stochastic model, we first set up a spatial domain Ω. We construct a domain with the appropriate surface, of which we impose zero flux boundary conditions.

In general the domain Ω is meshed into V tetrahedra shaped subvolumes $\Omega_k, k \in \{1,..,V\}$ such that,

$$\Omega = \bigcup_{k=1}^{V} \Omega_k, \quad \text{and} \quad \Omega_i \cap \Omega_j = \emptyset, \forall i \neq j, \quad i,j \in \{1,..,V\}.$$

The state of the system at time t is described as the matrix $X(t) \in \mathbb{Z}_+^{N \times V}$, where N is the number of species we are considering in the pathway and the components, X_{ik} are random variables, describing the molecular copy number of the ith species in Ω_k. Thus, the kth column of our state matrix, $X_{.k}$ describes the copy number of each species in Ω_k. Whereas, the ith row, $X_{i.}$, gives the copy number of species i in each subvolume.

3.3.1 Reaction events

First, we consider reactions that can occur due to molecular contact at correct orientation only. Later in the section we introduce diffusive jumps between subvolumes, as a simultaneous linear death process in one subvolume and a creation of that same species in a randomly chosen adjacent subvolume and also a convection term to model active transport along microtubules.

We assume that the species of our system, within each subvolume are uniformly distributed throughout that domain and in thermal equilibrium, such that the motion of each molecule is random. We can then consider the probability of a collision occuring between two reactant molecules. The fact that the copy number of mRNAs and proteins within cells is a small, nonnegative integer and that there is a probability that a certain reaction will occur when two molecules meet make a discrete, stochastic

description of signalling pathways necessary (Higham 2007).

Let the initial state of each subvolume be known at time $t = t_0$. The present state of the system can only make a jump to another state in the statespace, if they are connected via a transition vector $\mathbf{n} \in \{\mathbf{n_r}\}$, $r \in 1,..,M$, where $\mathbf{n_r}$ is the rth column of the stoichiometric matrix for the system, describing the change in copy number of the species involved in the allowed rth reaction and M is the total number of reactions involved. Two restrictions are imposed on the next possible reaction. One is that only a subset of the total number of reactions of our system is valid within each subvolume and hence within the compartment it lies. Condition two, is the avoidance of unfeasible reactions, ensuring all components of the state remain positive. Thus, if a single feasible reaction r is randomly selected, within a certain subvolume after an infinitesimal time dt, the corresponding state vector for that subvolume, $\mathbf{X}_{.k}(t_0)$ transforms to,

$$\mathbf{X}_{.k} \rightarrow \mathbf{X}'_{.k} = \mathbf{X}_{.k}(t_0 + dt) = \mathbf{X}_{.k}(t_0) + \mathbf{n_r}.$$

Let $P(X,t)$ be the probability of being in state X at time t. If we assume the Markovian property of conditional probability holds i.e. that the possible future state of a system is only dependent on the present state and not on any past states,

$$P(\mathbf{X}(t_n) = \mathbf{X_n}|\mathbf{X}(t_{n-1}) = \mathbf{X_{n-1}}|...|\mathbf{X}(t_0) = \mathbf{X_0}) = P(\mathbf{X}(t_n) = \mathbf{X_n}|\mathbf{X}(t_{n-1}) = \mathbf{X_{n-1}})$$

where $t_0 < t_1 < ... < t_{n-1} < t_n$.

Then, on the consideration of reactions alone, the temporal evolution of the probability distribution of each state in the statespace S is governed by the chemical master equation:

$$\frac{dP(X,t)}{dt} = \sum_{k=1}^{V}\sum_{r=1}^{M}\{\omega_r(\mathbf{X}_{.\mathbf{k}} - \mathbf{n_r})P(\mathbf{X}_{.\mathbf{1}},...,\mathbf{X}_{.\mathbf{k}} - \mathbf{n_r},..,\mathbf{X}_{.\mathbf{V}},t) - \omega_r(\mathbf{X}_{.\mathbf{k}})P(\mathbf{X},t)\} \tag{3.7}$$

$$= \mathbf{M}P(X,t) \tag{3.8}$$

where $\omega_r(\mathbf{X}_{.k}(t)) = \lim_{t\to 0}\frac{P(\mathbf{X}_{.k}(t)+n_r,t+dt)-P(\mathbf{X}_{.k}(t),t)}{dt}$, is known as the propensity function, describing the probability of making a transition between two connected states within the time interval $[t,t+dt]$.

The probability distribution governing the next time step, dt now follows an exponential distribution. We assume each of the reaction propensity functions follow mass action kinetics. For example, if reaction r^*, was to occur in Ω_k and was bimolecular, such that $X_{ik}+X_{jk}\rightarrow X_{lk}$, the corresponding propensity would be, $\omega_{r^*}(X_{.k}) = c_{r^*k}X_{ik}X_{jk}, \quad r^* \in r^c \subset \{1,..,M\}$

The propensity $\omega_r, r \in \{1,..,M\}$ has the same form for all subvolumes, k, whereas the reaction coefficient c_r scales with the volume $|\Omega_k|$ of the subvolume such that $c_{rk} = \hat{c_r}/|\Omega_k|$, where $\hat{c_r}$ is a constant.

3.3.2 Diffusion events

We now introduce diffusive jumps from one subvolume to a randomly selected adjacent subvolume as a first order event. In doing so we treat the diffusive process in a similar way to a reactive process and consider the probability of this transition taking place i.e. we consider the probability for one of the ith species to make a jump from the jth subvolume to an adjacent kth subvolume. Hence, we consider the following reaction and corresponding propensities,

$$x_{ij} \rightarrow x_{ik}, \quad a_i(X_{i.}) = d_{jik} x_{ij}.$$

Where each d is only non zero for connected mesh elements, $d_{j.j} = 0.$ and d_{jik} now depends on the volume of Ω_j, Ω_k and the macroscopic diffusion coefficient D. See Engblom et al. (2009) and Drawert et al. (2012) for more details.

For a system now accounting for reaction and diffusion, the temporal evolution of the probability distribution of each state in the statespace S is governed by the RDME,

$$\begin{aligned}\frac{dP(X,t)}{dt} &= \sum_{j=1}^{V}\sum_{k=1}^{V}\sum_{i=1}^{N}\{a_{jik}(X_{ij}-s_{jk})P(\mathbf{X_1}.,...,\mathbf{X_i}. - \mathbf{s_{jk}},...,\mathbf{X_N}.,t) - a_{jik}(X_{i.})P(\mathbf{X},t)\} \\ &+ \sum_{k=1}^{V}\sum_{r=1}^{M}\{\omega_{kr}(\mathbf{X}_{.\mathbf{k}} - \mathbf{n_{rk}})P(\mathbf{X}_{.\mathbf{1}},...,\mathbf{X}_{.\mathbf{k}} - \mathbf{n_{rk}},...,\mathbf{X}_{.\mathbf{V}},t) - \omega_{kr}(\mathbf{X}_{.\mathbf{k}})P(\mathbf{X},t)\} \\ &= \mathbf{D}P(X,t)\end{aligned}$$

Active transport events

We also consider active transport events in our model. In order to implement active transport along microtubules we first have to consider the microtubule density distribution within the cell; binding and unbinding of specific species to microtubule fibres and the velocity field that will drive the transport mechanism.

Let ρ_k be the mean density of microtubules within subvolume Ω_k and let x_{fk}, x_{bk} be a free and bound species respectively. Then we consider the following binding reaction of a particular free molecular species to a microtubule filament, with corresponding propensity,

$$x_{fk} \rightarrow x_{bk} \quad \eta(\rho_k, x_{fk}(t)) = \rho_k A_f x_{fk},$$

where A is the binding rate of species $x_{f.}$.

Once a species is bound to the microtubule it can either be transported or unbind from the filament.

The unbinding reaction propensity follows first order mass action kinetics:

$$x_{bk} \rightarrow x_{fk} \quad \omega(x_{bk}(t)) = A_o x_{bk}.$$

A species bound to a microtubule can be transported along it in much the same way as a diffusive jump event but this time, there is a directional bias, governed by the velocity field.

Let $\vec{x} = (x_1, x_2, x_3)$ be a point within the computational domain, given in Cartesian coordinates. Then we let $v(\vec{x})$ describe the velocity field at that position.

This Markov jump event has the following state change and propensity:

$$x_{bi} \rightarrow x_{bj} \quad c_b(x_{b.}) = c_{b.} x_{b.},$$

where $c_{b.}$ is composed of the speed of transport, the number of filaments between the two connecting subvolume faces and the length of that section of filament.

We now arrive at a more general transport operator:

$$\frac{\partial P(X,t)}{\partial t} = (\mathbf{M} + \mathbf{T}) P(X,t), \quad \mathbf{T} = \mathbf{D} + \mathbf{C}$$

We are now able to sample the RDME and produce realisations that are statistically exact.

Thus, in order to capture the intrinsic stochasticity of reactions, we model our system mathematically as a continuous time, discrete space, Markov chain, taking values within the state space $S \in \mathbb{Z}_{+}^{N \times V}$ defined on the domain Ω.

Chapter 4

Modelling Synthetic GRNs: The n-Gene Repressilator

4.1 The Hes1 GRN: The one gene repressilator

The transcription factor, Hes1, is the final protein product synthesised indirectly via the Hes1 gene. It is a member of the family of basic helix-loop-helix (bHLH) transcription factors: Hes1-7. The Hes family suppresses transcription of genes that require a bHLH protein for their transcription. Hes1 contains a helix interrupting structure that binds to the N-box promoter region rather than the enhancer box (E-box) of its target genes, resulting in the Hes1 protein also being a transcriptional repressor of itself.

Healthy embryonic development relies on the proper oscillatory expression of Hes1. For instance, in neural progenitor cells (NPCs), Hes1 represses the expression of proneural genes, such as Ascl1. Oscillatory expression of Hes1, in turn, inducing oscillatory expression of proneural genes such as Ascl1 by cyclic repression, is vital

for the cell cycle progression of NPCs and hence, the maintenance of NPCs. Persistent over-expression of Hes1 in NPCs, thus, can both inhibit NPCs proliferation and expression of genes needed for NPC differentiation giving rise to quiescent NPCs (Kageyama et al. 2015, Shimojo et al. 2008). Hes1 is seen to be negligible during NPC differentiation, thus allowing for the persistent expression of proneural genes.

It is thus apparent that the spatio-temporal dynamics of Hes1, has a high impact on which genes are expressed and hence, cell fate decision making.

Oscillation of Hes1 protein levels has been shown to be governed by a negative feedback loop (Hirata et al. 2002), where the Hes1 protein binds to its own gene, suppressing transcription of the corresponding Hes1 mRNA.

Hes1 mRNA is expressed with a basal transcription rate via the Hes1 gene. Once transcribed, Hes1 mRNA travels out of the nucleus into the cytoplasm via diffusion. Once it reaches the ribosomes Hes1 protein synthesis can begin. Newly synthesised Hes1 protein travels from the cytoplasm to the nucleus to its target genes, one of which is its own Hes1 gene. Hes1 protein binds to the N-box promoter of the Hes1 gene, inhibiting transcription of Hes1 mRNA. The decline in Hes1 mRNA leads to no new synthesis of Hes1 protein, thus, the current Hes1 protein will shortly degrade, freeing up the occupied promoter, allowing for the transcription of Hes1 mRNA to occur again. Thus, completing the cycle.

4.1.1 The PDE model

We now consider a reaction diffusion model of the Hes1 system, originally developed by Sturrock et al. (2011*b*) and which will form the basis for the subsequent development of our n-gene repressilators later in this chapter. The difference between the

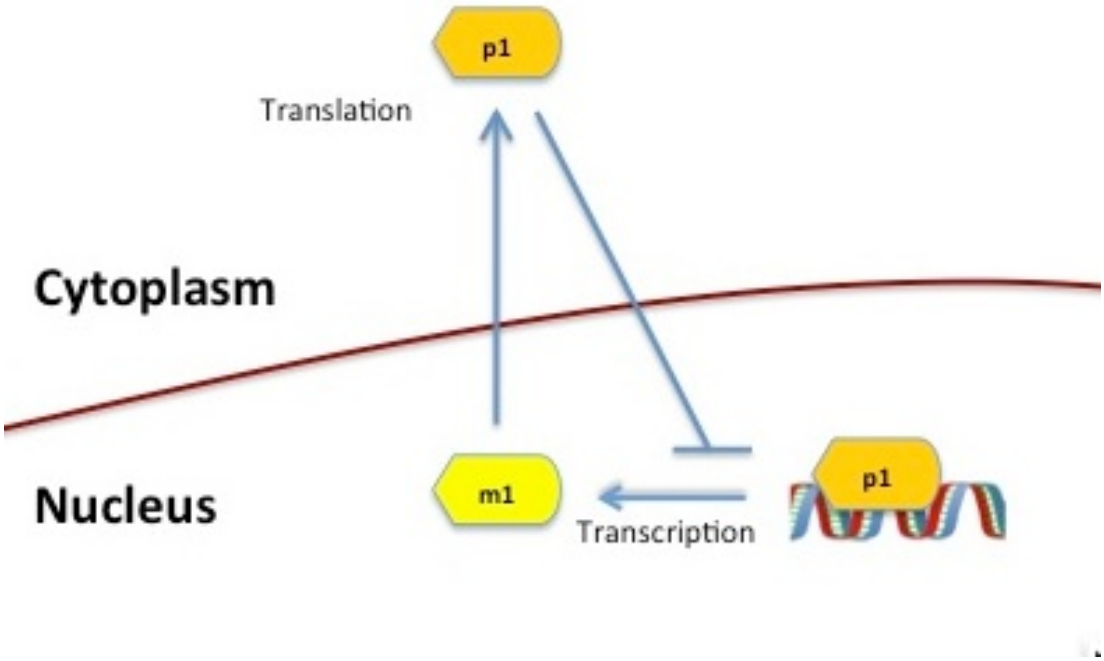

Figure 4.1: *A simple schematic diagram of the Hes1 GRN: The one-gene repressilator.*

model in this thesis and Sturrock et al. (2011*b*) is that we choose to specifically include gene sites and restrict transcription of mRNA to that region only. Sturrock et al. (2011*b*) allowed transcription to occur over the entire nucleus.

Let $m = m(x,y,t)$ be the concentration of hes1 mRNA and $p = p(x,y,t)$ be the concentration of Hes1 protein at position (x,y) at time t.

Let $(x_g, y_g) \in G$, where G is the set of all coordinates contained in the gene site; $(x_n, y_n) \in N$, where N is the set of all coordinates within the nucleus and $(x_c, y_c) \in C$, where C is the set of all possible coordinates within the (annular) cytoplasm.

We then model the Hes1 GRN with the following system of PDEs:

$$\begin{aligned} \frac{\partial m}{\partial t} &= D_m \nabla^2 m + \frac{\alpha_m}{1+p^h} H_g(x,y) - \delta_m m, \\ \frac{\partial p}{\partial t} &= D_p \nabla^2 p + \alpha_p m H_c(x,y) - \delta_p p, \end{aligned} \tag{4.1}$$

where α_m is the basal transcription rate, α_p is the protein synthesis rate, δ_m, δ_p are the mRNA and protein degradation rates, h is the hill function coefficient and D_m, D_p are the mRNA and protein diffusion coefficients. The two piecewise functions,

$$H_g(x,y) = \begin{cases} 1 & \text{if } (x,y) \in G, \\ 0 & \text{if } (x,y) \in C \cup N, \end{cases}$$

and

$$H_c(x,y) = \begin{cases} 1 & \text{if } (x,y) \in C, \\ 0 & \text{if } (x,y) \in G \cup N, \end{cases}$$

restrict the reactions of mRNA transcription to the Hes1 gene site location and protein translation to the cytoplasm. All other reactions of hes1 mRNA and Hes1 protein are global as are diffusive processes.

We apply the following initial conditions,

$$m(x,y,0) = p(x,y,0) = 0$$

and zero flux boundary conditions on the outside cellular membrane,

$$\frac{\partial m(\mathbf{x},t)}{\partial \mathbf{n}} = \frac{\partial p(\mathbf{x},t)}{\partial \mathbf{n}} = 0,$$

where $\mathbf{n}$ is the unit normal to the cellular membrane. Continuity conditions (of concentration and flux) are imposed on the inner nuclear membrane (described below).

Let us first consider the domain to consist of an outer annular cytoplasm, an inner (initially) annular nucleus and a nuclear-bound circular gene site centred at the origin. The cytoplasm is contained between the two most outer circles of radii $10\mu m$ and $3\mu m$ and the nucleus is contained between the two most inner circles with radii $3\mu m$ and $0.5\mu m$. The gene site is circular with a radius of $0.5\mu m$.

The parameter values and their description for the Hes1 system are given in Table 4.1 below.

Parameter	Reaction Description	Parameter value and unit
α_m	Transcription rate of mRNA	$1 min^{-1}$
α_p	Protein translation rate	$20 min^{-1}$
δ_m	Degradation rate of mRNA	$0.08 min^{-1}$
δ_p	Degradation rate of protein	$0.02 min^{-1}$
h	Hill function coefficient	5
$D_p = D_m = D$	Molecular diffusion	variable ($m^2 min^{-1}$)
r	Distance of gene site centre from the origin	variable (μm)

Table 4.1: *List of parameters and their values for the Hes1 reaction-diffusion model.*

The Hes1 PDE model was solved numerically using the multiphysics simulation software COMSOL 4.3a. The finite element method was implemented using triangular basis elements and Lagrange quadratic basis functions along with a backward Euler time stepping method of integration (and subsequently throughout the thesis). This method was chosen due to its ability to handle complicated geometries and boundaries with relative ease.

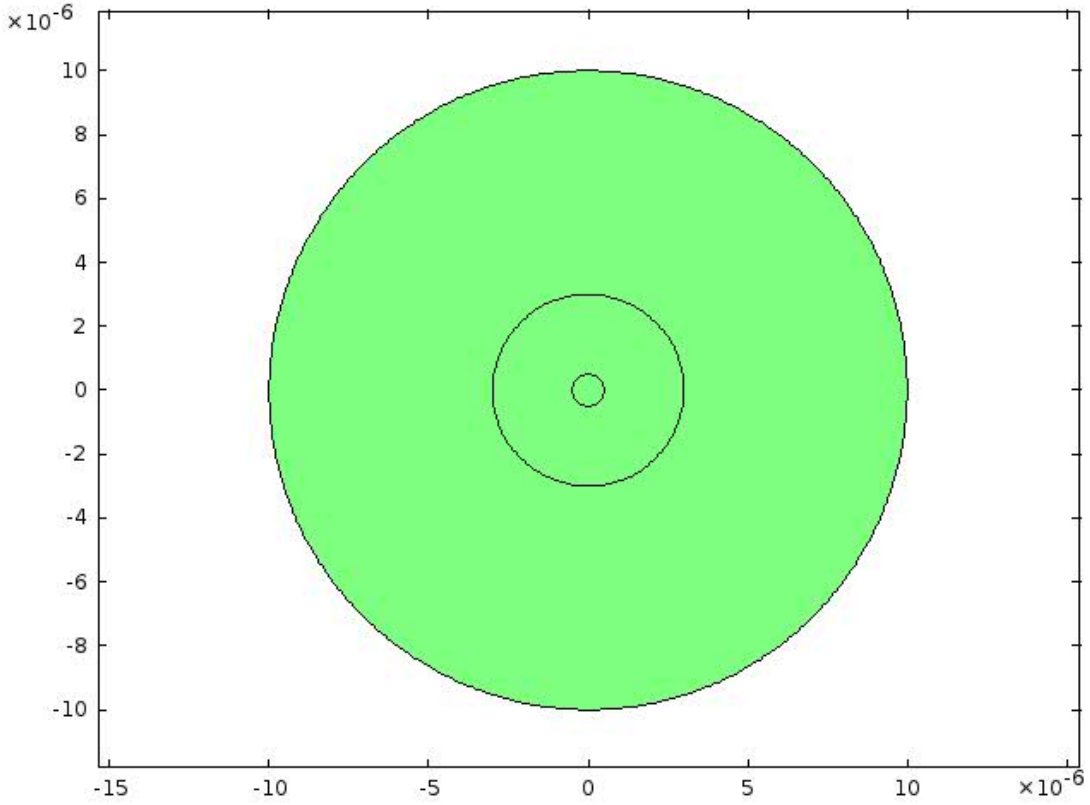

Figure 4.2: *Diagram of the computational spatial domain of the 2D circular cell: The cytoplasm is contained between the two most outer circles of radii* $10\mu m$ *and* $3\mu m$ *and the nucleus is contained between the two most inner circles with radii* $3\mu m$ *and* $0.5\mu m$*. The central gene site is circular with radius of* $0.5\mu m$*.*

The first set of simulations considered a gene site centred at the origin with a variable value of the diffusion coefficients. Depending on the specific diffusion parameter used, the dynamics of hes1 mRNA and Hes1 protein vary from steady state, to damped oscillations, to sustained oscillations. For sustained oscillations to be achieved, the diffusion parameter D, must fall within a certain range. Above and below this range, these oscillations are damped and as the parameter is further increased/decreased, steady states are achieved. As the diffusion coefficient is lowered, a steady state of zero (no expression: mRNA can not survive long enough to make it to the ribosomes for protein synthesis to occur) was achieved for protein and as diffusion coefficient increased

above this range the protein reached a positive, non zero steady state (persistent expression). If we now only consider the range of diffusion coefficients that allow for sustained oscillations (5×10^{-14} to 1×10^{-13}m^{2}min^{-1}), there is still yet a further interesting change in dynamics. Depending on the specific value within this range used, we can see a drastic change in the size of the amplitude of oscillation but we also see a change in the period of oscillation, which we know to have utmost importance in the next step of cell fate decision making.

The diffusion reference value, $D = 5 \times 10^{-14}$m^{2}min^{-1} was used in the following spatial profile. Steady, spatially-symmetrical, spatio-temporal oscillations were recorded.

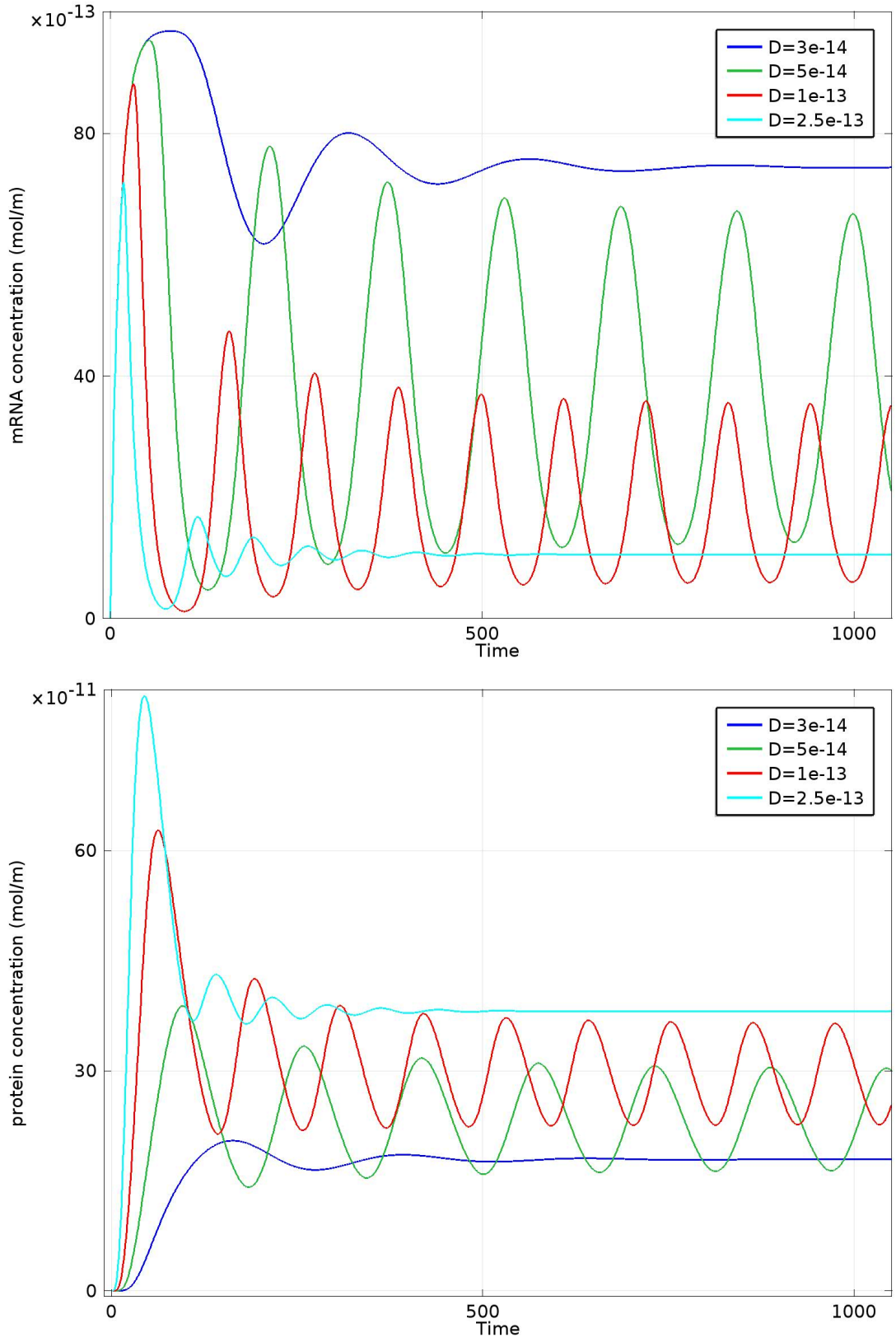

Figure 4.3: *One-gene repressilator: Plots showing the concentration levels of hes1 mRNA (top) and Hes1 protein (bottom) integrated over the entire domain against time. Gene site located centred at origin, variable diffusion coefficient D.*

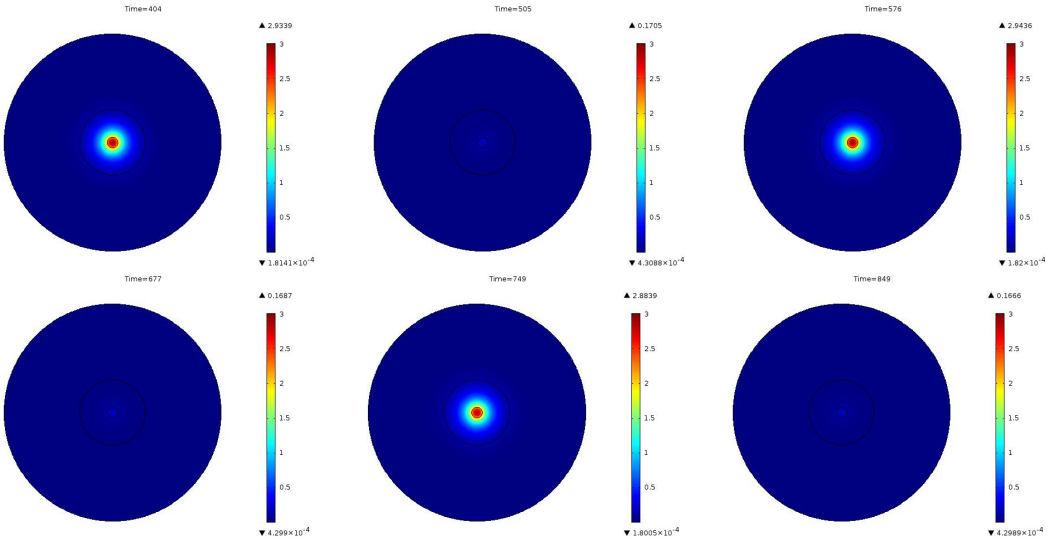

Figure 4.4: *Plots showing the spatio-temporal profiles of hes1 mRNA (the one gene repressilator) with the gene site centred at the origin.* $D = 5 \times 10^{-14}\text{m}^2\text{min}^{-1}$.

We now present computational simulations where we use the reference diffusion coefficient but vary the distance of the gene site from the origin. Where there were once oscillatory dynamics for the origin-located gene first turns to damped oscillations as the distance is increased and then for a certain threshold distance, these oscillations are totally lost as the gene location is placed closer to the nuclear membrane. Thus, a similar pattern appears to emerge as we vary the position of the gene site within the nucleus as occurs when varying the diffusion coefficient (Sturrock et al. 2011*b*). There seems to be a very intrinsic relationship between the diffusion parameter and the radial spatial parameter of the gene location. As the gene location gets further away from the centre of the nucleus, the diffusion parameter has to be decreased in order to sustain spatial-temporal oscillatory dynamics.

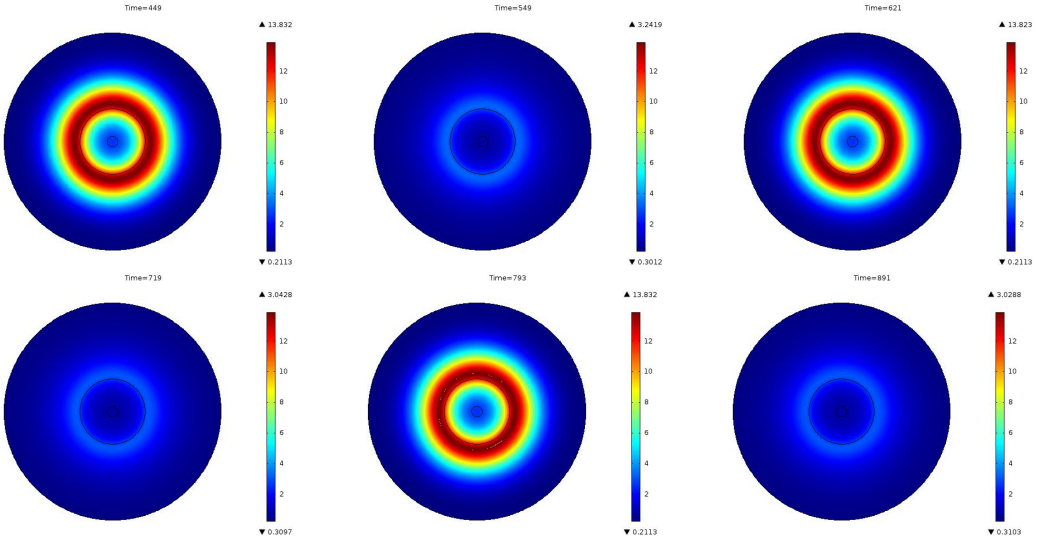

Figure 4.5: *Plots showing the spatio-temporal profiles of Hes1 protein (the one gene repressilator) with the gene site centred at the origin.* $D = 5 \times 10^{-14}\text{m}^2\text{min}^{-1}$.

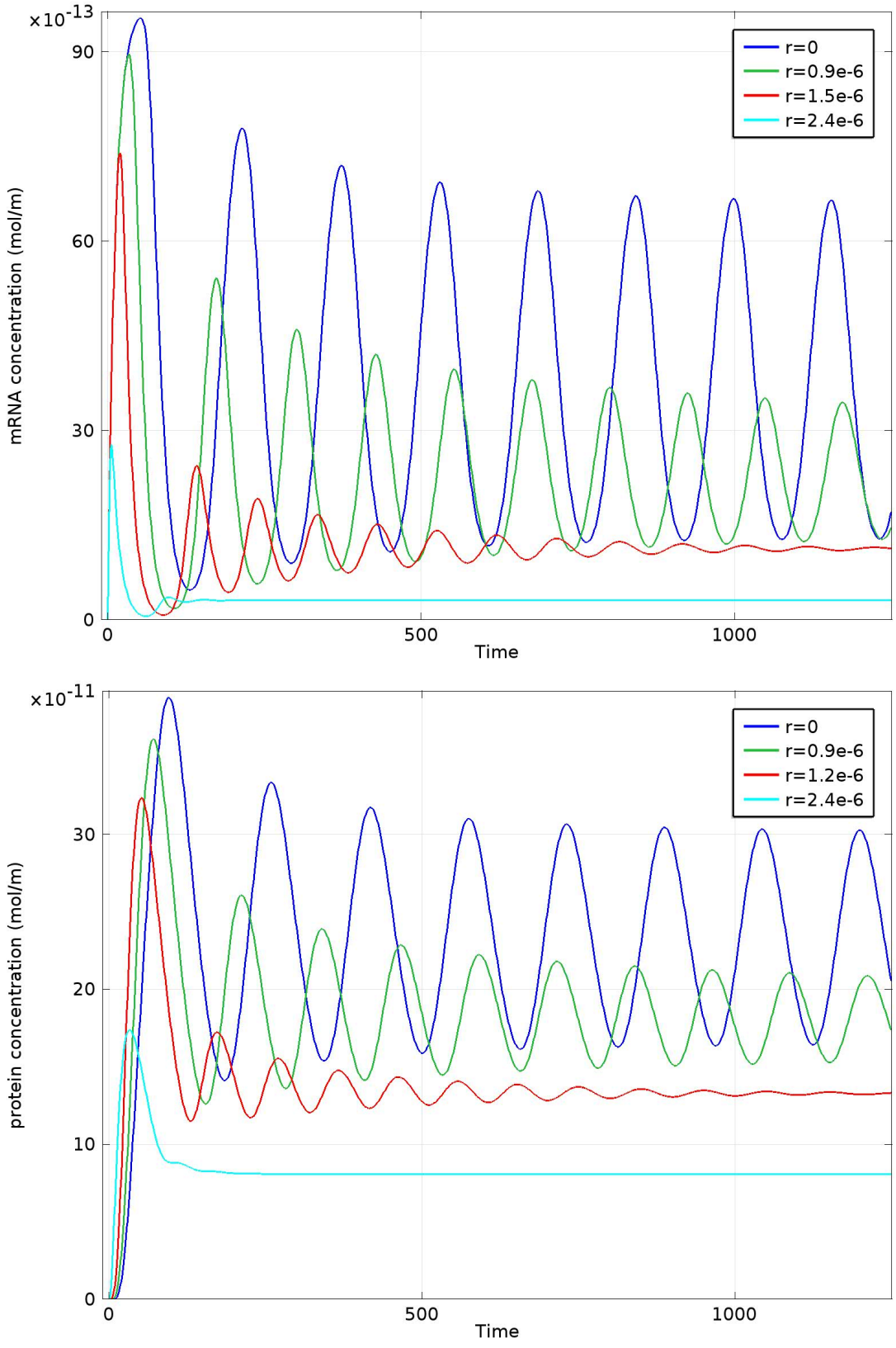

Figure 4.6: *One-gene repressilator: Plots showing the concentration levels of (top) hes1 mRNA and (bottom) Hes1 protein integrated over the entire domain against time. Varying the distance of the gene site from the origin:* $D = 5 \times 10^{-14}\mathrm{m}^2\mathrm{min}^{-1}$.

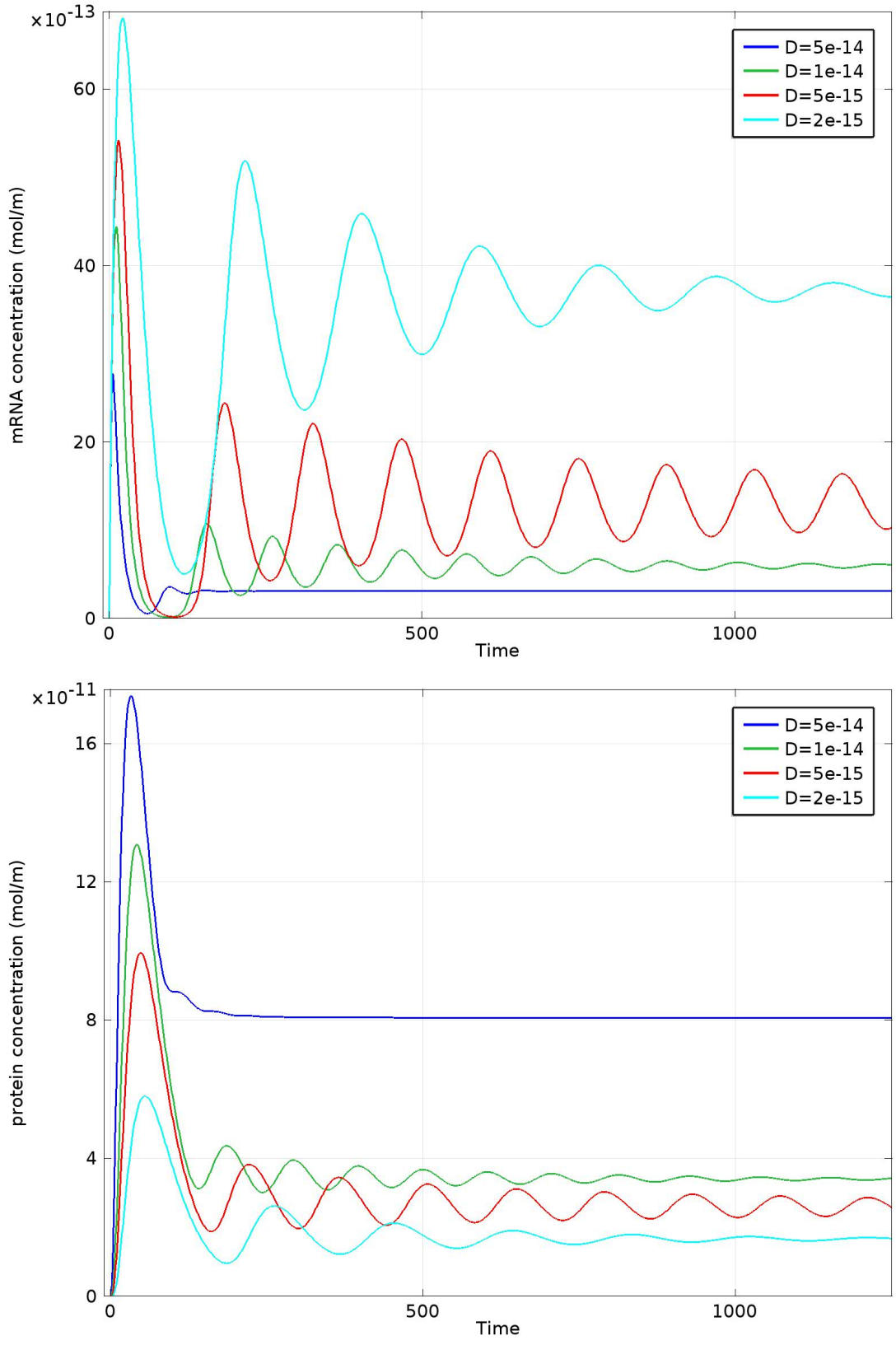

Figure 4.7: *One-gene repressilator: Plots showing the concentration levels of hes1 mRNA (top) and Hes1 protein (bottom) integrated over the entire domain against time. Gene site close to the nuclear membrane, fixed distance* $r = 2.4\mu m$*, - varying D.*

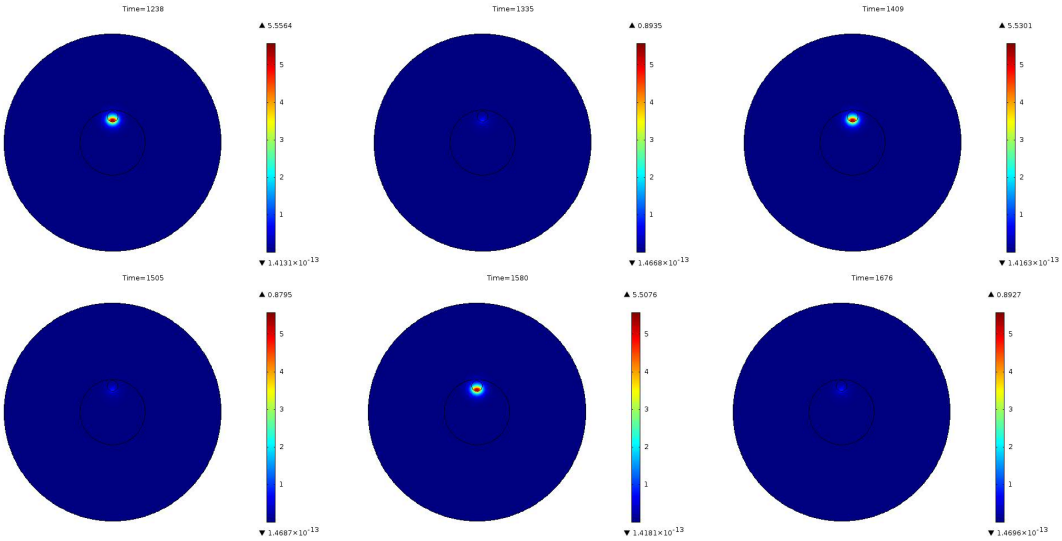

Figure 4.8: *Plots showing the spatio-temporal profiles of hes1 mRNA with the gene site close to the nuclear membrane.*

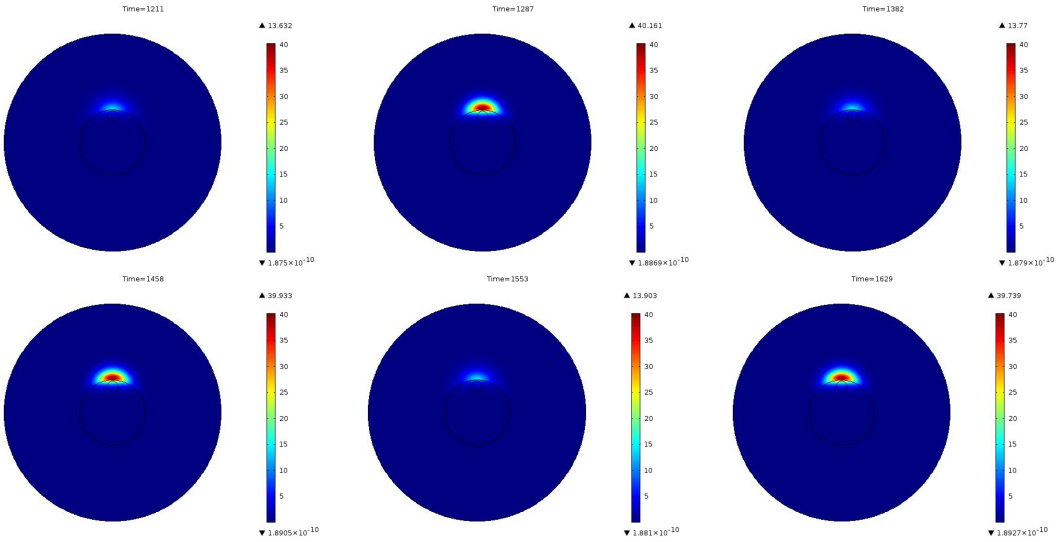

Figure 4.9: *Plots showing the spatio-temporal profiles of Hes1 protein with the gene site close to the nuclear membrane.*

4.2 Model Extension: The n-gene Repressilator

The Hes1 GRN is an example of a naturally occurring, auto-inhibiting GRN, and one can "classify" the Hes1 GRN as a one-gene repressilator. We now discuss a theoretical extension of the Hes1 system to a class of GRNs, termed n-gene repressilators. The n-gene repressilator ($n > 1$) has the same topological structure as a ring network, where there are $i = 1, \ldots, n$ genes and their products with each gene in the network being connected to the previous and subsequent gene. This means that the output of a node i is fed into the next node $i+1$ in the ring, and that the input to node i was from the output of node $i-1$. The repressive nature of the ring specifically means that each node (gene) is suppressed by the node before it, and in turn that node suppresses the following neighbour node in the ring. The structure of a 5-gene repressilator is shown in Figure 4.10.

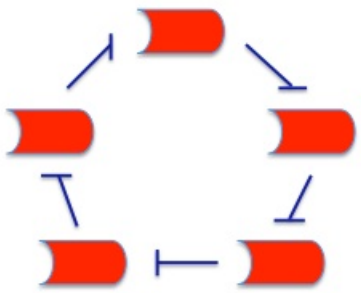

Figure 4.10: *Schematic diagram of the 5-gene repressilator, $n = 5$. Each node is suppressed by the preceding node within the ring and in turn suppresses the output of the succeeding node.*

n-gene repressilators were originally conceived as theoretical intracellular systems in the field of Systems Biology (and now Synthetic Biology). A genetic toggle switch was constructed and mathematically modelled and tested by Gardner et al. (2000) who believed that switches like this are important in development. A simple ODE model

was presented and conditions for bistability were investigated. The switch was constructed with two repressible promoters arranged in a mutually inhibitory circuit. Two steady states were seen to be possible, one low (or zero: the repressed state) and the other high (or non zero: the active state) with a switch in concentration happening when an inducing stimulus of the repressed state was added. Here, we model a similar system, referring to it as the two-gene repressilator.

Elowitz & Leibler (2000) set out to build a synthetic GRN as an artificial biological clock. The result is now what is known as the original repressilator: three-gene repressilator. They wanted to gain more insight into the design behind natural biological clocks such as the those responsible for circadian rhythms.

To build their clock they began with three different repressors and their corresponding promoters. The repressor proteins were based around the TetR, LacI, and λcI system, and their corresponding promoters were, respectfully, pTet, pLac, and λPR. pTet is repressed by TetR, pLac (from the lac operon of lactose metabolism) is repressed by LacI, and λPR is repressed by cI. The mathematical model developed to study this system was an ODE system, using a hill function for mRNA synthesis and mass action kinetics for molecular degradation and protein synthesis. A second model used the discrete stochastic Markov chain algorithm, known commonly as the Gillespie algorithm. Temporal oscillatory dynamics were achieved. Since then, others have also sought to model repressilator systems (Muller et al. 2006, Oliveira et al. 2015).

To date, the majority of models of synthetic GRNs do not consider the geometry of the cell or localisation of biological events. Some temporal models have been extended to account for compartmentalisation, considering nuclear cytoplasmic ratios but this still can not take into account that mRNA synthesis does not happen everywhere in the nucleus but at exact gene-site locations within the nucleus. ODE models also do not take into account the spatial distribution of organelles and genes within each compartment.

Thus, the role that symmetry or asymmetry of organelle and gene site location within the cell may play is completely ignored.

In this chapter we investigate how the location of gene sites within the nucleus and the diffusion coefficients of molecules can heavily affect the ultimate spatio-temporal dynamics of synthetic GRNs, specifically n-gene repressilators.

The next step in gene expression is heavily governed by how genetic information is processed. This information can be stored in the spatio-temporal dynamics of the gene regulatory networks. Specific genes can induce or suppress hundreds of other genes. It has been found that some genes have an early response to GRN dynamics while others may be affected by the same pathway later on (Ashall et al. 2009, Nelson et al. 2004). This of course may be due to post-translational modifications, with particular modification dependent upon temporal stimulant profile and cross talk. However, here we have decided to focus on how the spatial organisation of the cell may affect these short- and long-term dynamics. The amplitude and most importantly, period of oscillations have vital implications for subsequent gene expression. For instance, as a cell becomes cancerous or progresses in malignancy, the shape of the cell may become distorted. Work by Meyers et al. (2006) and Neves et al. (2008) demonstrate how cellular size and shape control spatial information flow and hence, GRN activity.

We now introduce the general n-gene repressilator as a system of reaction diffusion equations, (n$\geq$2):

For $i = 2, .., n$,

$$\frac{\partial m_i}{\partial t} = D_{m_i}\nabla^2 m_i + H_{g_i}\frac{\alpha_{m_i}}{1+p_{i-1}^h} - \delta_{m_i} m_i, \tag{4.2}$$

$$\frac{\partial p_i}{\partial t} = D_{p_i}\nabla^2 p_i + H_c\alpha_{p_i} m_i - \delta_{p_i} p_i, \tag{4.3}$$

and for $i = 1$,

$$\frac{dm_1}{dt} = D_{m_1}\nabla^2 m_1 + H_{g_1}\frac{\alpha_{m_1}}{1+p_n^h} - \delta_{m_1} m_1 \tag{4.4}$$

$$\frac{dp_1}{dt} = D_{p_1}\nabla^2 p_1 + H_c\alpha_{p_1} m_1 - \delta_{p_1} p_1, \tag{4.5}$$

where,

$$H_{g_i}(x,y) = \begin{cases} 1, & \text{if } (x,y) \in G_i, \\ 0, & \text{if } (x,y) \in C \cup N \cup_{j\neq i}^{n} G_j, \end{cases}$$

and

$$H_c(x,y) = \begin{cases} 1, & \text{if } (x,y) \in C, \\ 0, & \text{if } (x,y) \in \cup_{i=1}^{n} G_i \cup N \end{cases}$$

are the two step functions restricting the reactions of mRNA transcription and protein translation to the appropriate cellular compartments. All other reactions and diffusion of mRNA and protein species are global. The following Table 4.2 states the parameter values used.

Parameter	Reaction Description	Parameter value and unit
α_{m_i}	Transcription rate of mRNA	$1 min^{-1}$
α_{p_i}	Protein translation rate	$20 min^{-1}$
δ_{m_i}	Degradation rate of mRNA	$0.08 min^{-1}$
δ_{p_i}	Degradation rate of protein	$0.02 min^{-1}$
h	Hill function coefficient	5
$D_{p_i} = D_{m_i} = D$	Molecular diffusion	variable ($m^2 min^{-1}$)
r	Distance of gene site centre from the origin	variable (μm)

Table 4.2: *List of parameters and their values for the n-gene repressilator PDE model*

The n-gene repressilator model is solved with the following initial conditions,

$$m_i(x,y,0) = p_i(x,y,0) = 0$$

subject to zero-flux boundary conditions on the cellular membrane,

$$\frac{\partial m_i(\mathbf{x},t)}{\partial \mathbf{n}} = \frac{\partial p_i(\mathbf{x},t)}{\partial \mathbf{n}} = 0,$$

where $\mathbf{n}$ is the unit normal to the cellular membrane and continuity conditions across the nuclear membrane (concentration and flux).

4.3 The two-gene repressilator

We now consider the two-gene repressilator reaction-diffusion model.

Let $m_1 = m_1(x,y,t)$, $m_2 = m_2(x,y,t)$, $p_1 = p_1(x,y,t)$ and $p_2 = p_2(x,y,t)$ be mRNA1, mRNA2, protein1 and protein2 concentrations respectively. We define the reaction diffusion system representing this GRN as the following set of equations.

$$\begin{aligned}
\frac{\partial m_1}{\partial t} &= D_{m_1}\nabla^2 m_1 + H_{g1}\frac{\alpha_{m_1}}{1+p_2^h} - \delta_{m_1} m_1, \\
\frac{\partial p_1}{\partial t} &= D_{p_1}\nabla^2 p_1 + H_c \alpha_{p_1} m_1 - \delta_{p_1} p_1, \\
\frac{\partial m_2}{\partial t} &= D_{m_2}\nabla^2 m_2 + H_{g2}\frac{\alpha_{m_2}}{1+p_1^h} - \delta_{m_2} m_2, \\
\frac{\partial p_2}{\partial t} &= D_{p_2}\nabla^2 p_2 + H_c \alpha_{p_2} m_2 - \delta_{p_2} p_2.
\end{aligned} \tag{4.6}$$

Where,

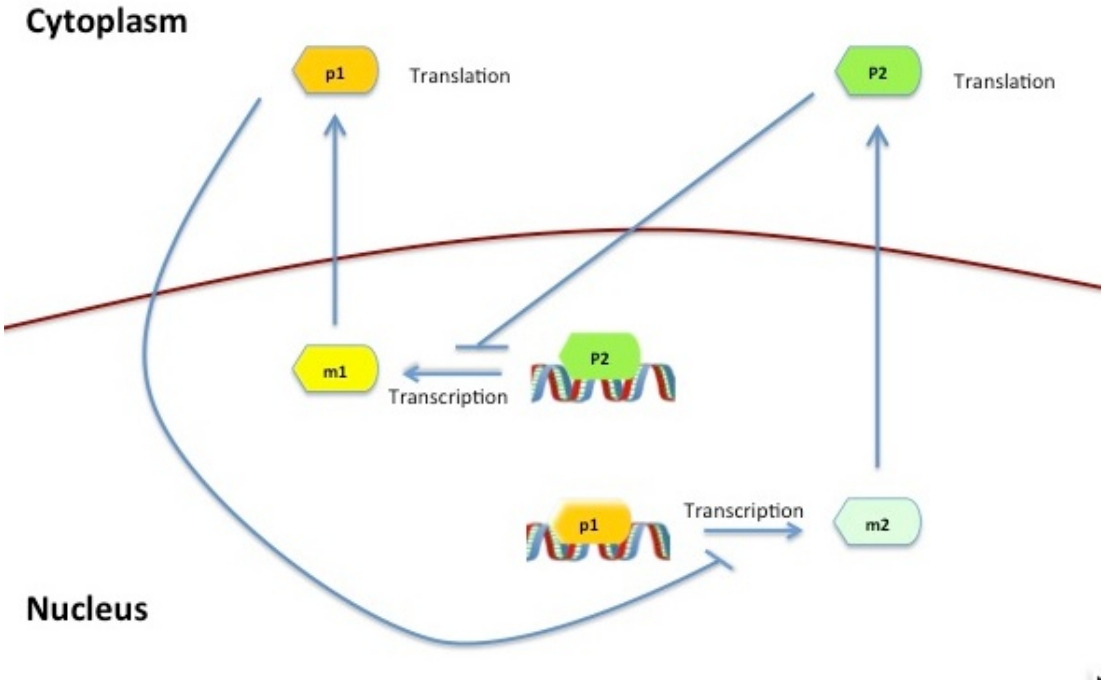

Figure 4.11: *Schematic diagram of the two-gene repressilator, $n = 2$.*

$$H_{g_1}(x,y) = \begin{cases} 1 & \text{if } (x,y) \in G_i \\ 0 & \text{if } (x,y) \in C \cup N \cup G_2, \end{cases}$$

$$H_{g_2}(x,y) = \begin{cases} 1 & \text{if } (x,y) \in G_i \\ 0 & \text{if } (x,y) \in C \cup N \cup G_1 \end{cases}$$

and

$$H_c(x,y) = \begin{cases} 1 & \text{if } (x,y) \in C \\ 0 & \text{if } (x,y) \in G_1 \cup G_2 \cup N \end{cases}$$

To be solved with the following initial conditions,

$$m_1(x,y,0) = m_2(x,y,0) = p_1(x,y,0) = p_2(x,y,0) = 0$$

subject to the following boundary conditions,

$$\frac{\partial m_1(\mathbf{x},t)}{\partial \mathbf{n}} = \frac{\partial m_2(\mathbf{x},t)}{\partial \mathbf{n}} = \frac{\partial p_1(\mathbf{x},t)}{\partial \mathbf{n}} = \frac{\partial p_2(\mathbf{x},t)}{\partial \mathbf{n}} = 0,$$

where $\mathbf{n}$ is the unit normal to the cellular membrane and $\mathbf{x}$ represents all of the coordinates on the cellular membrane.

With the initial conditions of zero i.e initially zero protein and mRNA concentration, this means that both mRNA can be maximally transcribed with rate α_m.

In the first set of simulations, we shall consider a single shared gene site, hence, $G_1 = G_2$. Separate gene sites on non-intersecting spatial locations will be considered shortly, particularly how this adaptation affects the spatio-temporal dynamics of the GRN.

4.3.1 Cellular domain with a single shared gene site

For the two gene repressilator GRN PDE model with one single shared gene site, the geometry used is exactly the same as that used for the Hes1 PDE system discussed earlier. For the above stated parameter values and a diffusion coefficient of $D = 5 \times 10^{-14} m^2 min^{-1}$ and a single shared gene site centred at the origin, ($r = 0$), the following spatial and temporal profiles were observed, (see figures 4.12 for the temporal profile and, figures 4.13 and 4.14 for some spatial snapshots).

We found that, for the given parameter values, oscillations were very persistent until $t = 3500$ and then drastically transformed to steady state solution. This is quite distinctly different from ODE model predictions of the two-gene repressilator as they might contain a transient spike but no level of transient oscillations are seen. However, the asymptotic long term, steady state solution remains the same for both the ODE models and this PDE model. Although, due to the fact that we use identical parameters

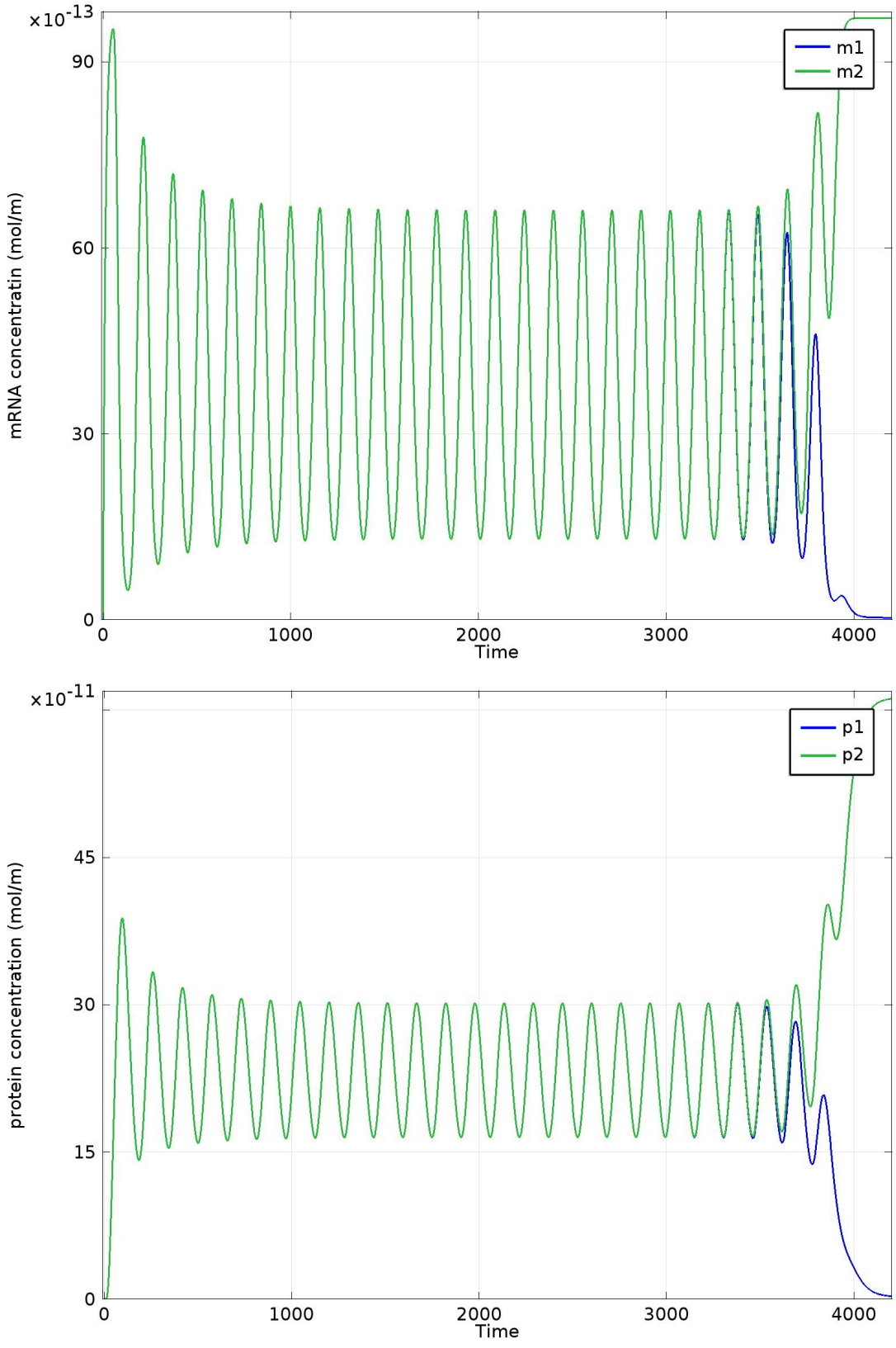

Figure 4.12: *Two-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. A shared gene site at the origin,* $D = 5 \times 10^{-14} m^2 min^{-1}$.

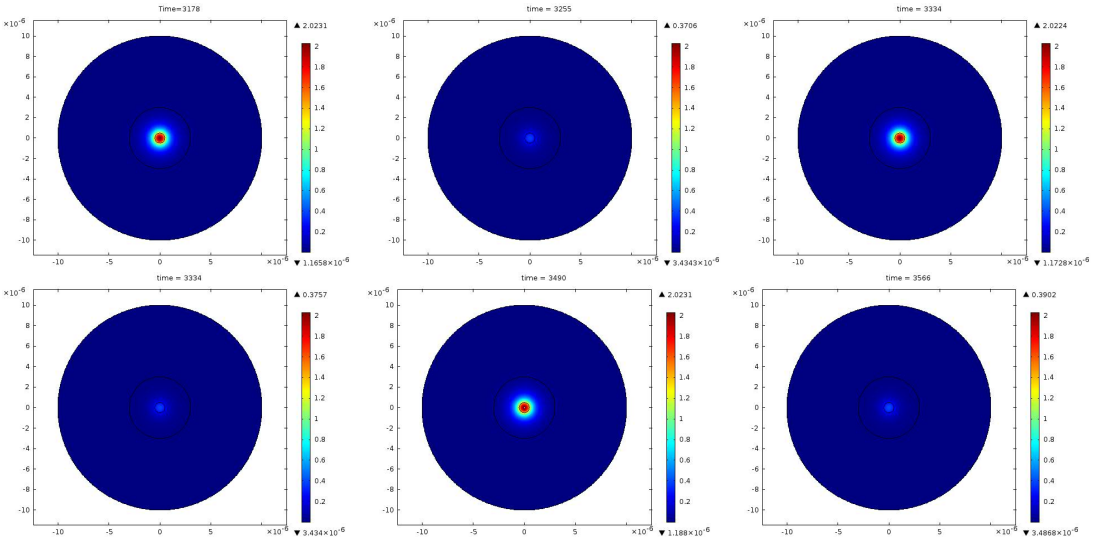

Figure 4.13: *Plots showing the spatio-temporal profiles of mRNA for the two gene repressilator with a shared gene site at the origin:* $D = 5 \times 10^{-14} m^2 min^{-1}$.

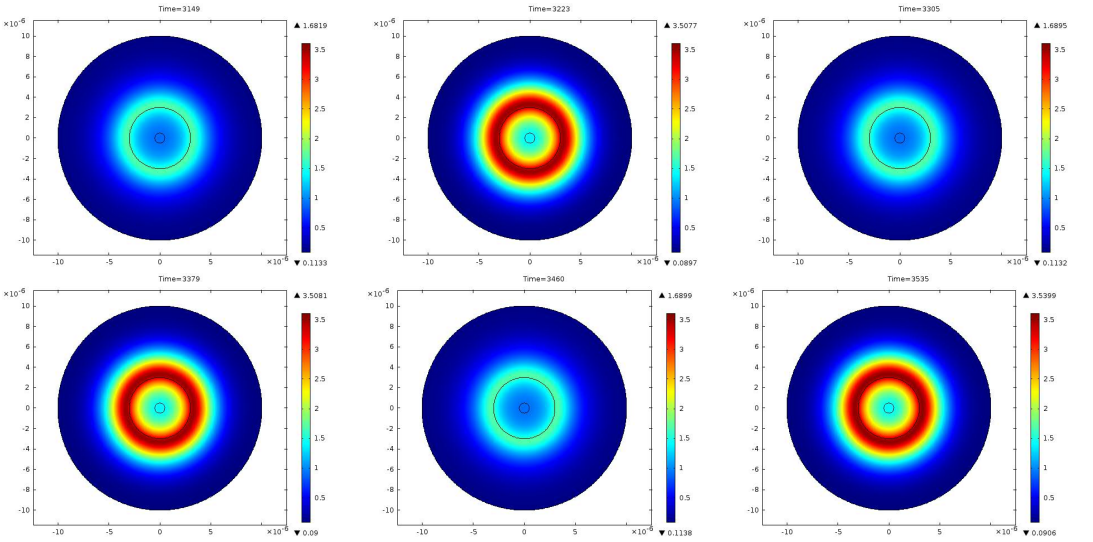

Figure 4.14: *Plots showing the spatio-temporal profiles of protein for the two gene repressilator with a shared gene site at the origin:* $D = 5 \times 10^{-14} m^2 min^{-1}$.

and initial conditions as in the one-gene repressilator model, this would make the solution of the $n = 1$ model a solution of this model. The transient presented here indicates that this solution is unstable. We attempted to investigate if the solution is dependent on mesh size and we find that for a decrease in maximum element size of 8×10^{-8} to 4×10^{-8} there is a tiny increase in the transient time of oscillation, see figure 4.15 to see how the mRNA (concentration integrated over the gene site) differs.

We tried to decrease the mesh size further but unfortunately, our computer did not have the memory capacity to do so. Thus, it is uncertain how long the transient time exists for, if it does at all, as we need a computer with greater capacity.

However, the ODE analysis done by Strelkowa & Barahona (2010) and Strelkowa & Barahona (2011) do find oscillatory transients in their n-gene repressilator models (but not for $n = 2$, only for $n \geq 3$). They propose that the transient time of quasi-stable oscillations increases with n (for even n) of which the first case they see is $n = 4$. This paper supports that this model needs further investigation.

Regardless of the minimal computing power that we have, we pressed on to investigate how a perturbation to the location of gene site one would impact our model. We translate gene site one $0.1\mu m$ along the positive x-axis, thus creating two gene sites, although overlapping. The solution showed a small transient, somewhat oscillatory phase and then a steady state solution was reached, (see figure 4.16).

We have introduced asymmetry into the system. mRNA1 now has a bias to reach the cytoplasm first and hence, for protein one to be synthesised first. Thus giving protein one the upper hand in becoming the dominant protein.

We then looked at how the transient time changed with a perturbation to the position of gene site one, (see figure 4.17). Comsol was very restrictive in how small it allowed us to make this perturbation but we were able to do increments of $0.01\mu m$. The transient

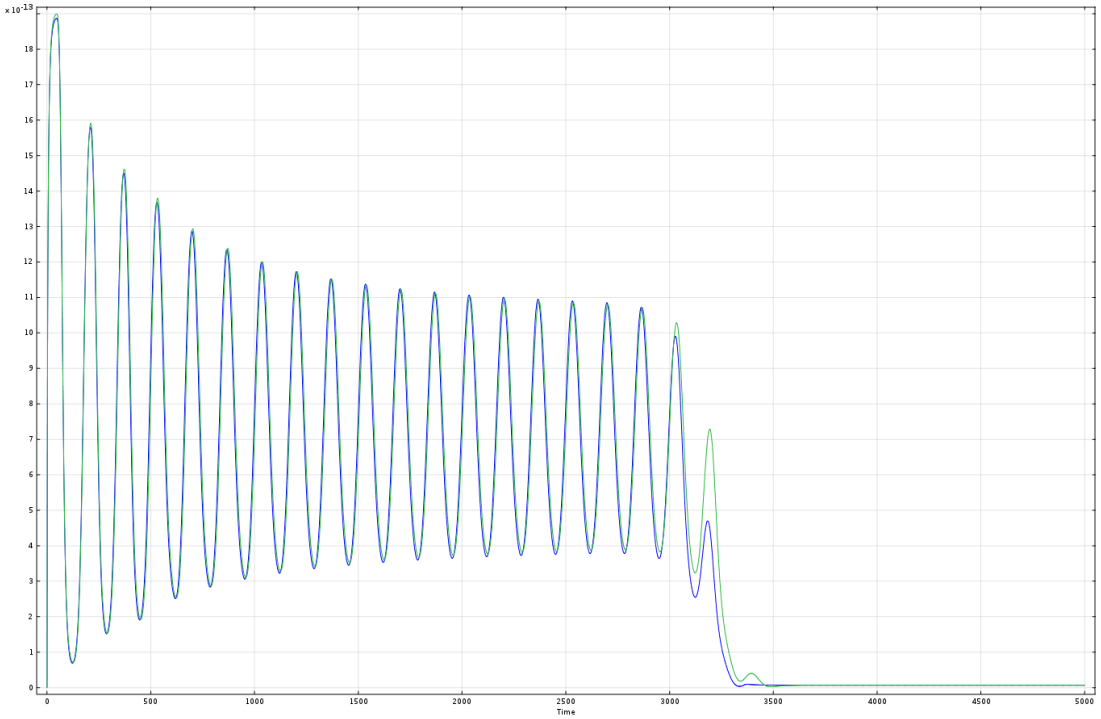

Figure 4.15: *Two-gene repressilator: Plot of mRNA1 concentration, (integration of concentration levels over the entire cell) against time. A shared gene site at the origin,* $D = 5 \times 10^{-14} m^2 min^{-1}$. *Blue is for a mesh size,* $h = 8 \times 10^{-8}$ *and green is* $h = 4 \times 10^{-8}$.

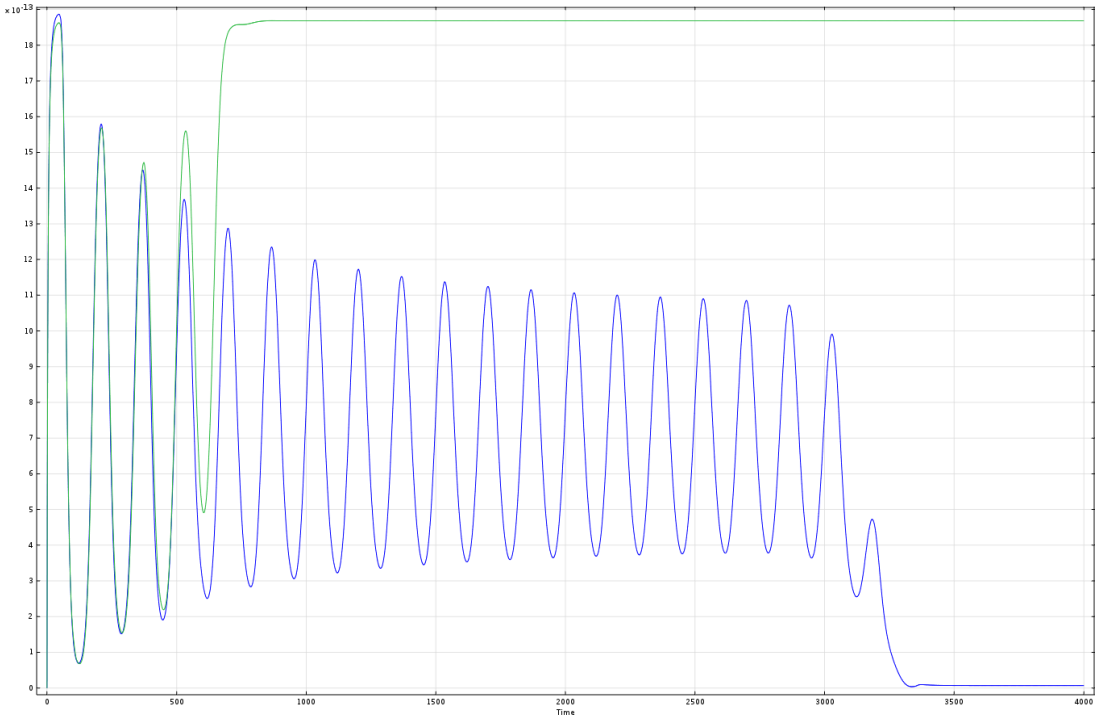

Figure 4.16: *Plot of mRNA1 concentrations, (integrated over gene site one) against time. Blue is for the case of a single shared gene site at the origin. Green is for the case where there are two overlapping gene sites: gene site one is slightly perturbed in its position from the origin, (at position $g_1 = (0.1\mu m, 0\mu m)$), while gene site two remains fixed at the origin, ($g_2 = (0\mu m, 0\mu m)$).*

time of the oscillations seen decreased very quickly.

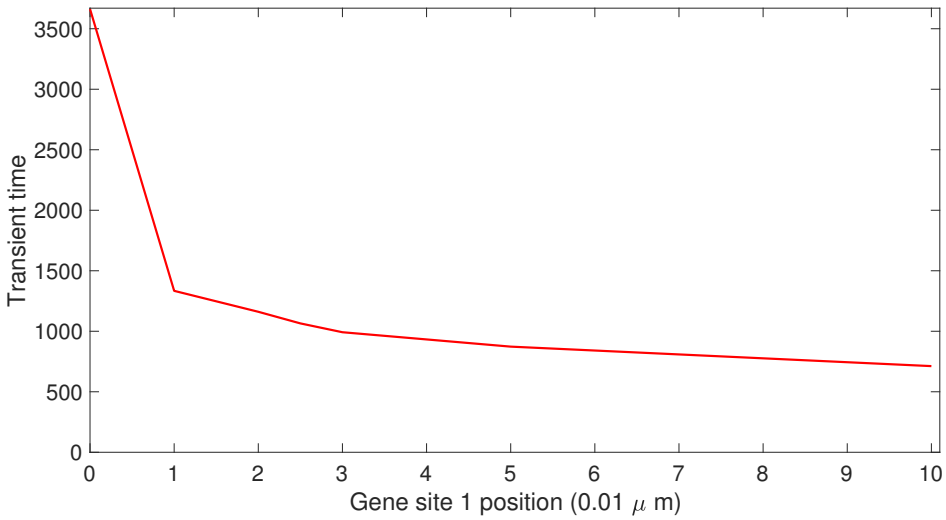

Figure 4.17: *Plot of transient time of oscillatory behaviour for the two-gene repressilator against the perturbation of the location of gene site one along the positive direction of the x-axis, (gene site two remaining fixed at the origin).*

We now consider how variation in the diffusion coefficient affects spatial-temporal dynamics for the two-gene repressilator with a single shared gene.

For the range of diffusion coefficients that give sustained oscillations for a significant amount of time, namely $D = 5 \times 10^{-14} - 1 \times 10^{-13} m^2 min^{-1}$ the period varied from 161 to 110 minutes; the amplitude of oscillation doubled and the transient time in dynamics varied from 2000 to 3500. Hence, again we see how the diffusion coefficient: a spatial parameter, is key in producing the spatio-temporal oscillations that are observed, agreeing further with Chaplain et al. (2014). As the diffusion coefficient is increased/decreased above/below this range, oscillations become heavily damped, see figure 4.18. As the diffusion coefficient is further increased/decreased again still, steady state solutions are reached. This, however, may not be too surprising if we consider that for large D the PDE system should act like the ODE system. In the ODE system there is only a transient spike before steady states are reached.

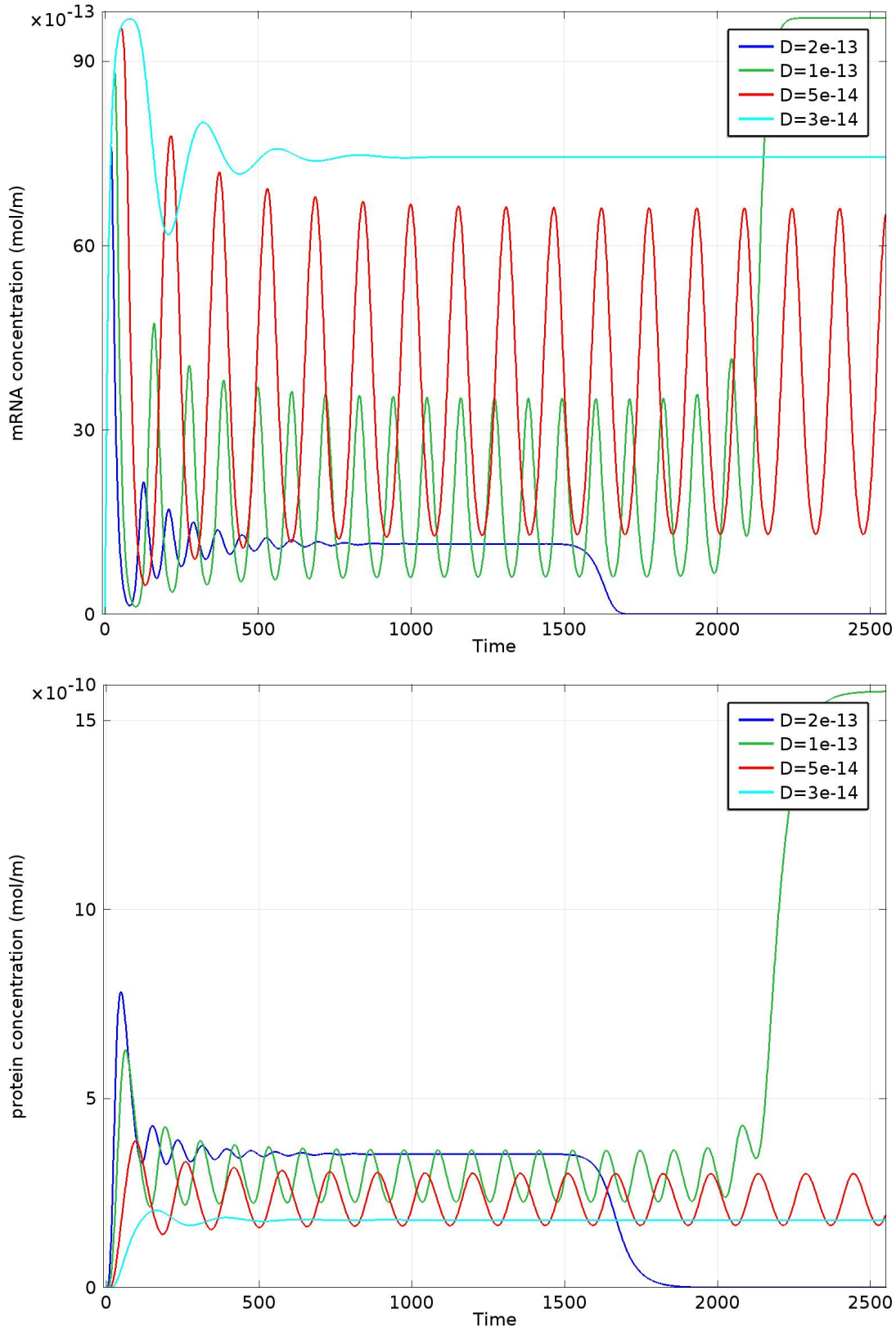

Figure 4.18: *Two-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. A shared gene site at the origin: variable D.*

In consideration of the major effect the diffusion coefficient had on the spatial-temporal dynamics we also choose to investigate the variation of the position of the centre of the gene site, r, within the nucleus i.e we vary the position of the localisation of mRNA transcription.

While doing so, we choose to keep the diffusion coefficient fixed at $D = 5 \times 10^{-14}$. As the centre of the gene site is moved further away from the origin, subsequently breaking the spatial symmetry of our domain, we see the same effect that decreasing the diffusion coefficient had on the spatial-temporal dynamics of the two gene repressilator with origin centred gene site. Thus, the amplitude of oscillation decreases by about half, oscillations die about 1200 minutes earlier and the period of oscillation decreases. Oscillations are completely damped at the half way mark i.e. between nuclear origin and nuclear membrane: $r = 1.5\mu m$ and after this, we see a steady state reached almost instantly.

We now consider varying the diffusion parameter for our model with a single gene site, but now with position close to the nuclear membrane i.e. $r = 2.4\mu m$. We find that all dynamics have either damped oscillations or are steady state solutions 4.20.

However, mRNA1 and mRNA2 are indeed transcribed at separate gene sites. Since we are investigating how the spatial aspects of our model affect spatial-temporal dynamics we must also consider separate gene sites and the affect this has on the spatio-temporal dynamics.

Previous papers, using ODE models, have recorded that the two gene repressilator exists in a bistable equilibrium state (O'Brien et al. 2012). However, in this PDE model, two different phases of dynamics are observed. The first phase consists of synchronised transient oscillations. This first phase lasts for a significant period of time, provided we remain within a range of symmetrical parameter values, equality in species' initial conditions and a symmetrical cellular structural arrangement.

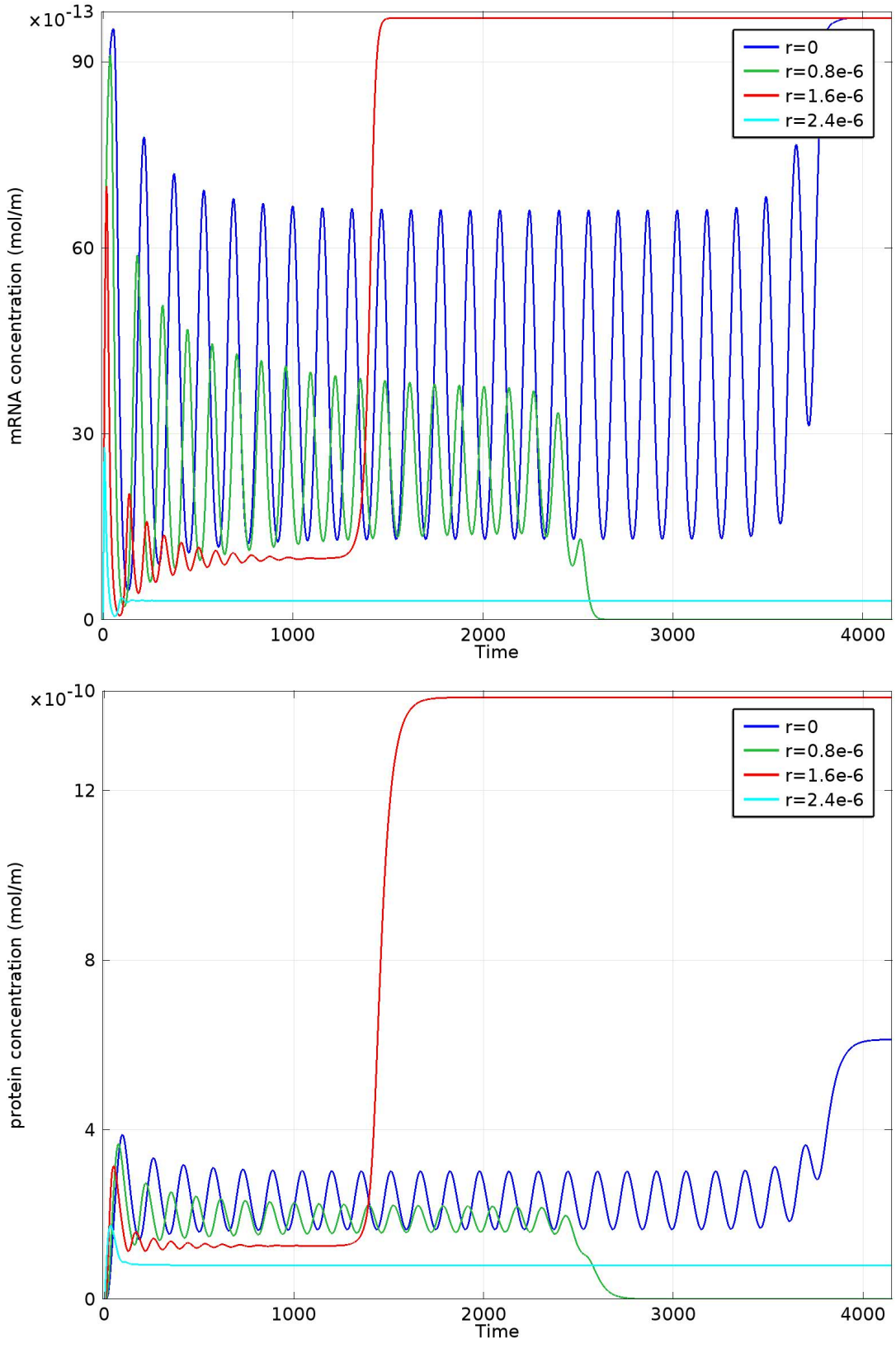

Figure 4.19: *Plots showing the concentration levels of mRNA and protein for the two gene repressilator over time for varying r, the position of the shared gene site.* $D = 5 \times 10^{-14} m^2 min^{-1}$. *Blue is* $r = 0\mu m$, *green is* $r = 0.8\mu m$, *red is* $r = 1.6\mu m$ *and cyan is* $r = 2.4\mu m$.

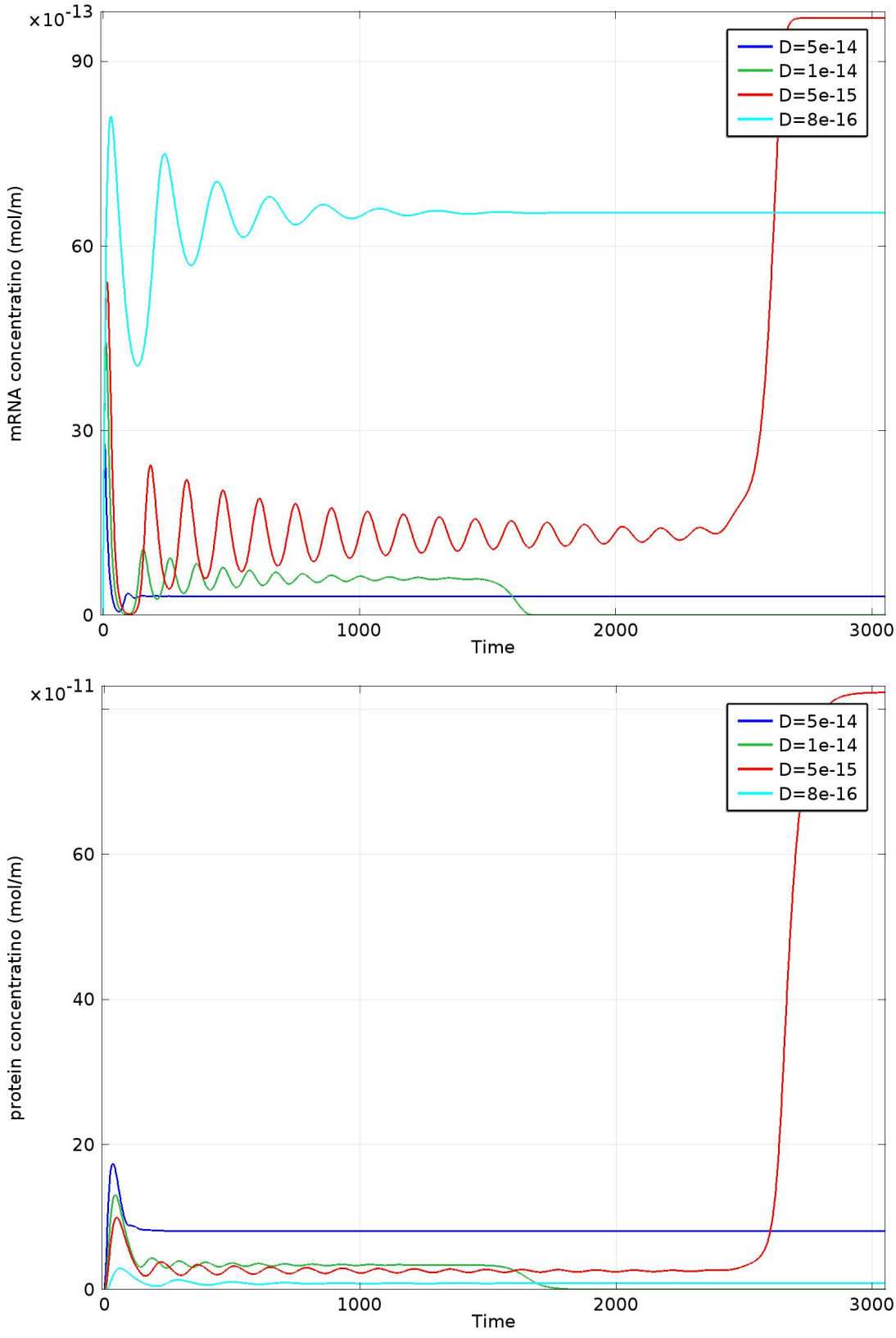

Figure 4.20: *Plots showing the concentration levels of mRNA and protein for the two gene repressilator over time with a shared gene site close to nuclear membrane ($r = 2.4\mu m$) and variable D. Blue is $D = 5 \times 10^{-14} m^2 min^{-1}$, green is $D = 1 \times 10^{-14} m^2 min^{-1}$, red is $D = 5 \times 10^{-15} m^2 min^{-1}$ and cyan is $D = 1 \times 10^{-15} m^2 min^{-1}$.*

These two different early and later term phases may bring some insight to the timing of early and late gene activation.

If there is asymmetry in the initial conditions, geometry or parameter values, then one protein will have a head start in its expression levels over the others. Once this happens, oscillations, if any, are then strongly damped and the protein with the head start soon reaches a maximum steady state and the other protein concentrations go down to zero.

4.3.2 Cellular domain with separate individual gene sites

We now consider how spatial-temporal dynamics of the two gene repressilator changes when we add in a second gene site to more faithfully represent the two distinct gene sites that will exist in reality.

Circular geometry with two separate gene sites

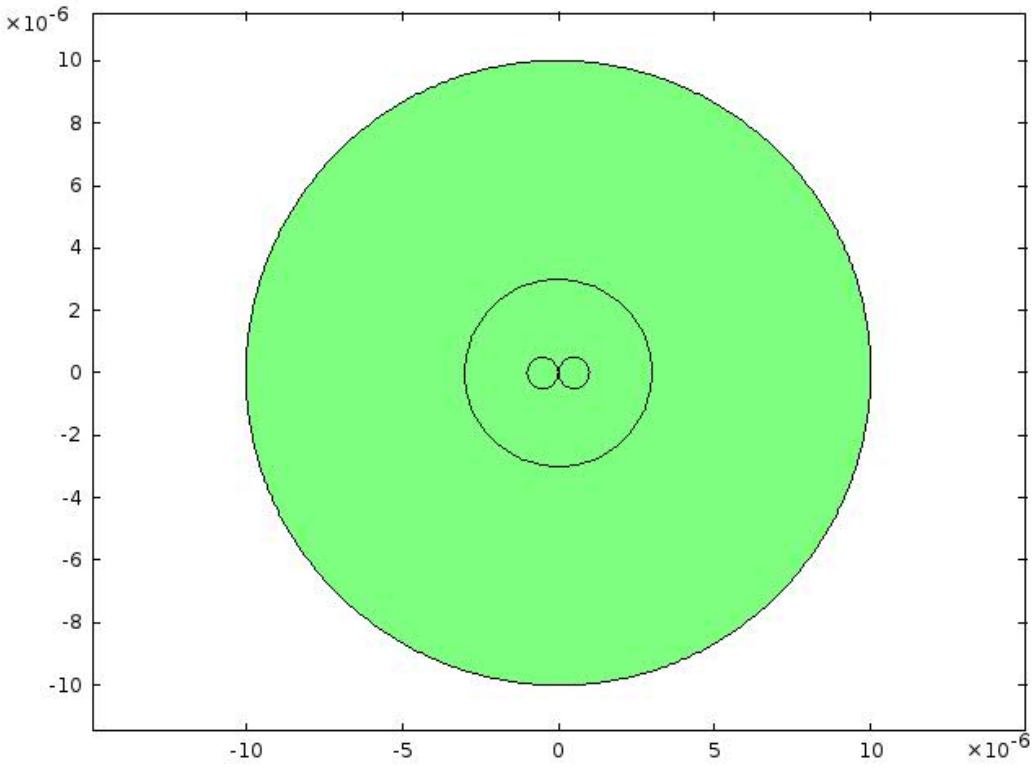

Figure 4.21: *Diagram of the computational spatial domain in the case of two separate gene sites. The cytoplasm is contained between two outer circles of radii* $10\mu m$ *and* $3\mu m$ *and the nucleus is the inner circle of radius* $3\mu m$*. Within the nucleus are the two separate gene sites, at coordinates* $(0.5\mu m, 0)$ *and* $(-0.5\mu m, 0)$*, each of radii* $0.5\mu m$*.*

Let each gene site be a non-intersecting circle of radius 0.5μm. We initially give the two separate gene sites coordinates $(0.5\mu m, 0)$ and $(-0.5\mu m, 0)$, the closest possible coordinates to the origin. The transcription of mRNA1 and mRNA2 are now confined to occur solely on gene site 1 and 2 respectively. Two phases of dynamics were again observed. For the specific parameters chosen, (same parameters as the shared gene

site, with the optimal value of diffusion being, $D = 5 \times 10^{-14}$), the initial transient dynamics were spatial-temporal quasi-stable oscillations. The difference in the period and amplitude of oscillations between the origin located shared gene site and origin located, neighbouring separate gene site model is relatively small. However, the change in the transient time of quasi-stable oscillations is significantly different. The transient oscillations start to die off after 1000 minutes after which the second phase of steady states begins. Thus, the change in dynamics occurred at a third of the time it took the system to do so with a single shared gene site, (see figure 4.22). However, once again, we can not be sure whether our model is independent of the mesh, so this might change with as the mesh is refined further.

In creating individual gene sites, an asymmetrical localisation of protein and mRNA is observed, (see figures 4.24 and 4.25). This could be an indicator as to why cellular polarisation can occur, which is of importance in embryonic development.

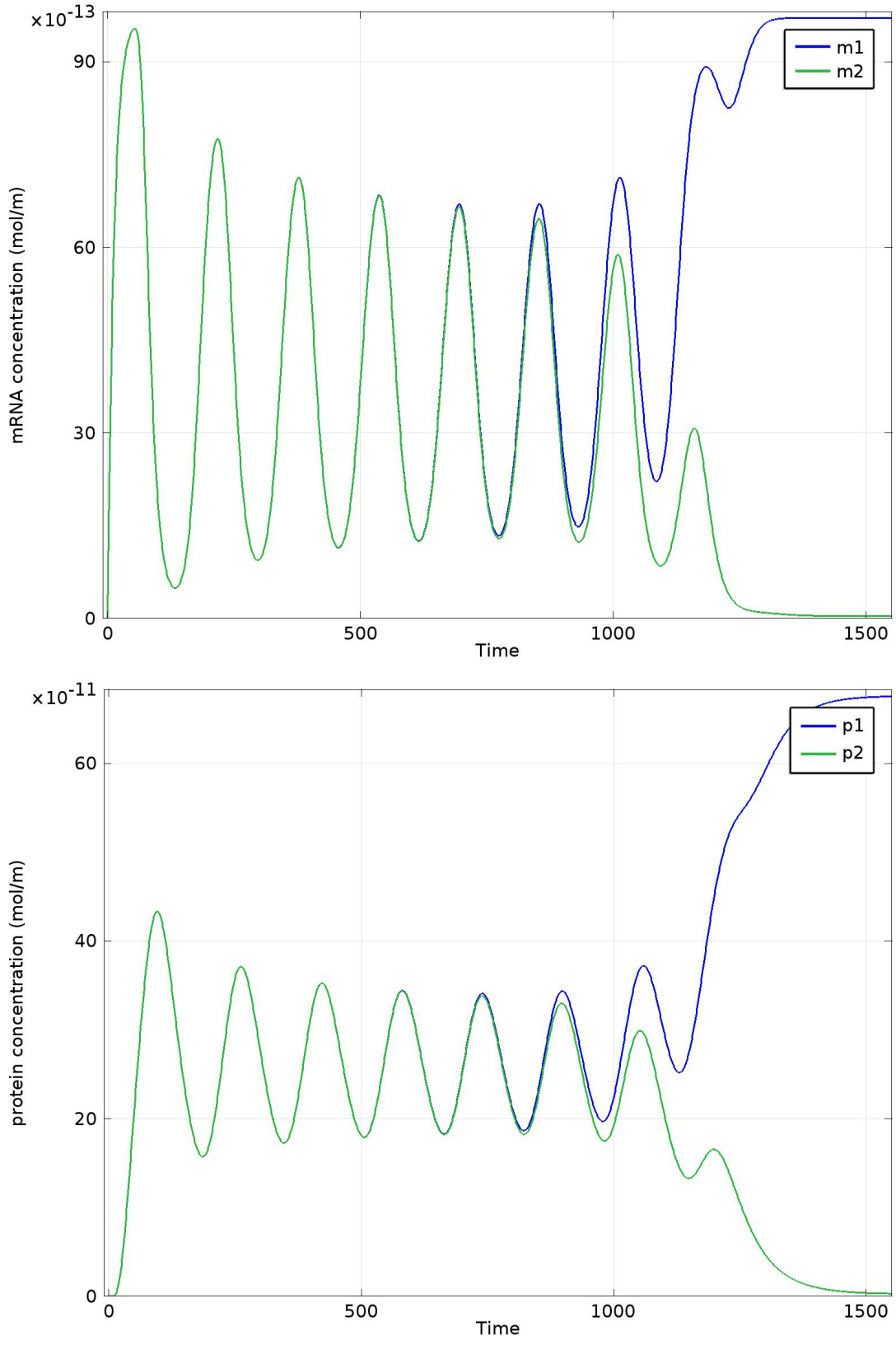

Figure 4.22: *The two gene repressilator: plots are of the integrated concentrations of mRNA (top) and protein (bottom) over the entire cell against time. Two separate but neighbouring gene sites at the origin. blue m1/p1 and Green m2/p2*

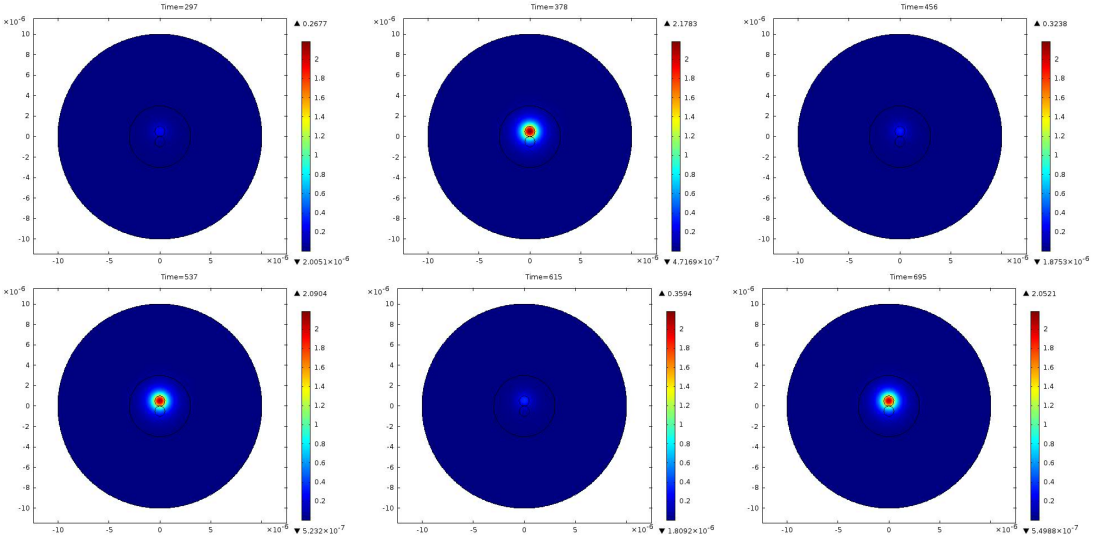

Figure 4.23: *The two gene repressilator: two separate but neighbouring gene sites at the origin - mRNA1. Spatial profile of two gene repressilator with two separate gene sites:* $D = 5 \times 10^{-14}$

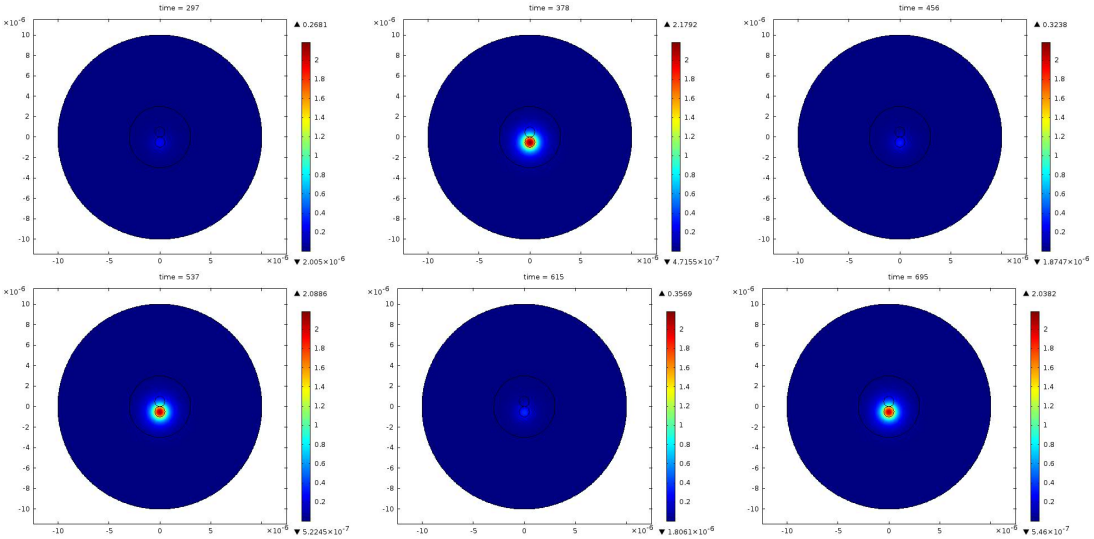

Figure 4.24: *The two gene repressilator: two separate but neighbouring gene sites at the origin - mRNA2. Spatial profile of two gene repressilator with two separate gene sites:* $D = 5 \times 10^{-14}$

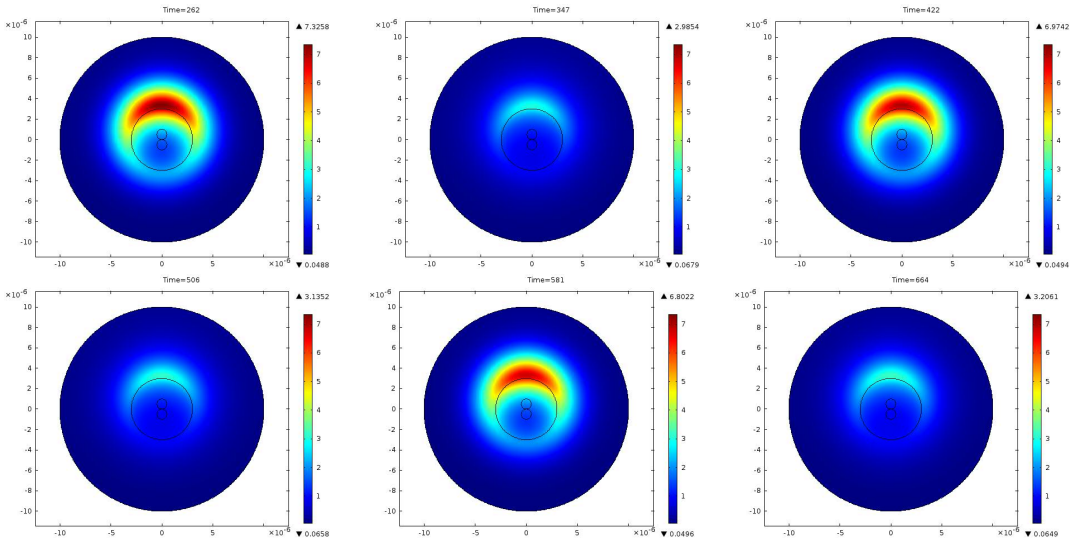

Figure 4.25: *The two gene repressilator: two separate but neighbouring gene sites at the origin - protein1. Spatial profile of two gene repressilator with two separate gene sites:* $D = 5 \times 10^{-14}$

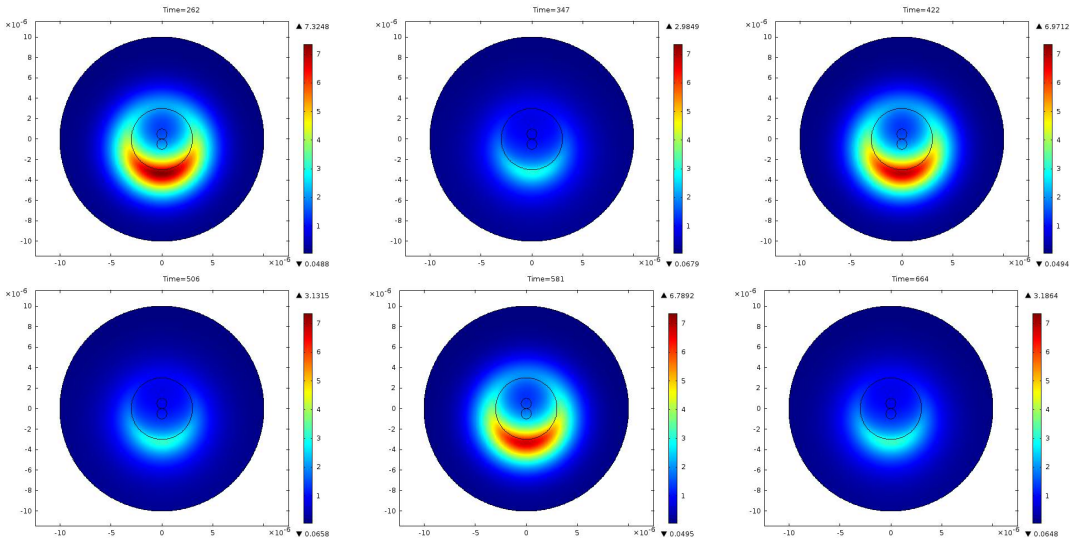

Figure 4.26: *The two gene repressilator: two separate but neighbouring gene sites at the origin - protein2. Spatial profile of two gene repressilator with two separate gene sites:* $D = 5 \times 10^{-14}$

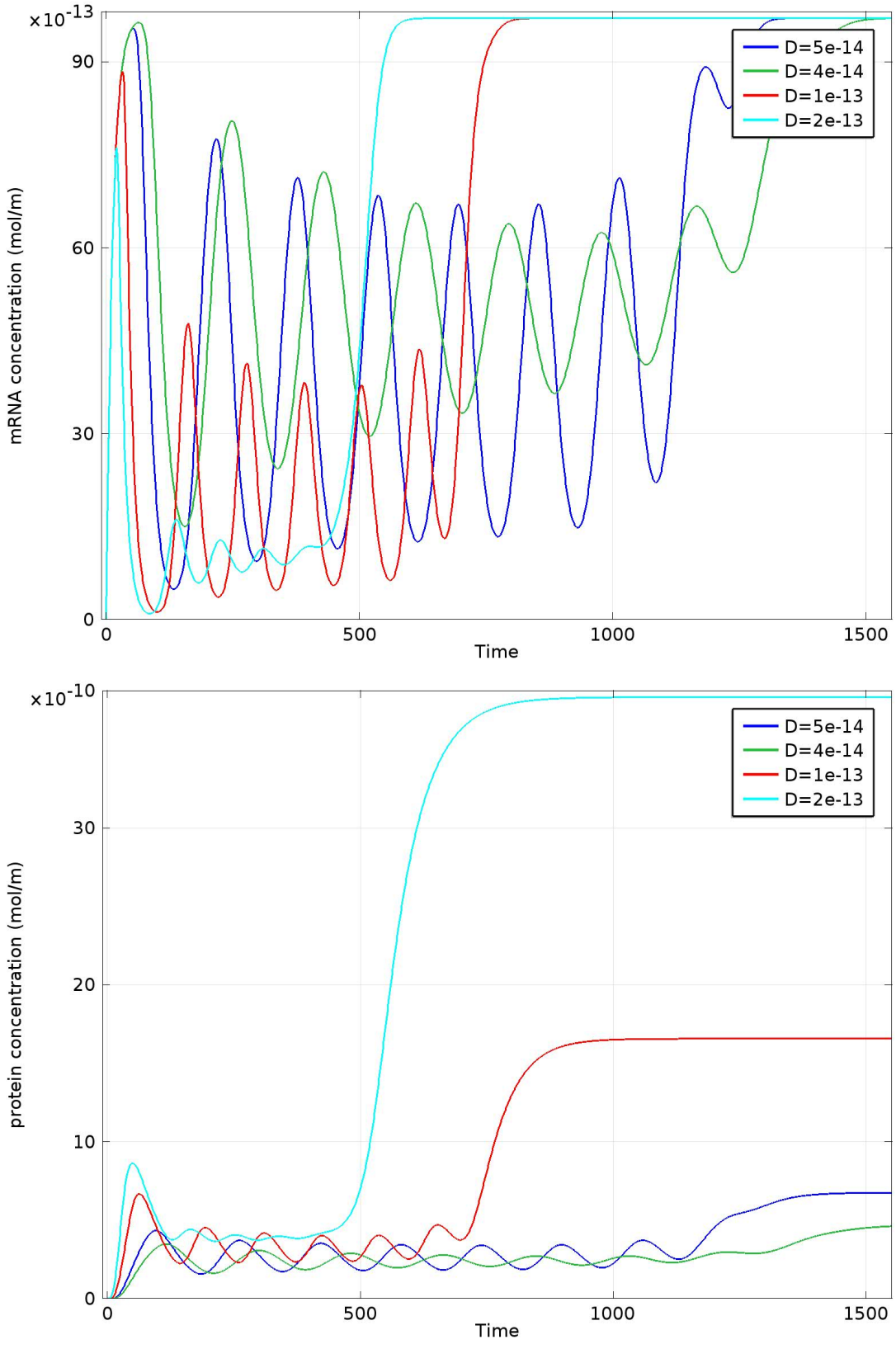

Figure 4.27: *Two-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Two separate but neighbouring gene sites close to the origin: Varying D.* $g_1 = (0.5\mu m, 0), g_2 = (-0.5\mu m, 0)$ *:- blue* $D = 5 \times 10^{-14}$, *Green* $D = 4 \times 10^{-14}$ *and Cyan* $D = 1 \times 10^{-13}$

In creating two separate gene sites, symmetry of the domain was broken. However, there is still a symmetry intact due to the fact that each gene site is the same distance away from the cytoplasm and in turn each repressing protein synthesised in the cytoplasm is an equal distance away from its target gene as the other.

Let us now consider two scenarios of varying gene site location. First we will keep symmetry in both the x-axis and y-axis i.e. we will vary the position of the gene1 and gene2 sites, each along the positive and negative x-axis respectively i.e the gene sites have coordinates, $(x+0.5\mu m, 0)$ and $(-x-0.5\mu m, 0)$ respectively. Second we will consider breaking symmetry in one of these axis and move the gene sites together cojointly as neighbours i.e. the gene sites have coordinates $(0.5\mu m, y)$ and $(-0.5\mu m, y)$ respectively.

In the first case, we see that the transient time of the initial phase of dynamics is relatively unaffected as x is increased, (see figure 4.28). Although, we do note that as the gene sites are moved further apart the amplitude of the oscillations of the protein species particularly quite drastically increased while the period only ever so slightly increases. In consideration of varying the diffusion parameter, we see that there are two phases for a specific range but that within this range the timing of phase change is quite distinctly slower for a higher diffusion coefficient, (see figure 4.29).

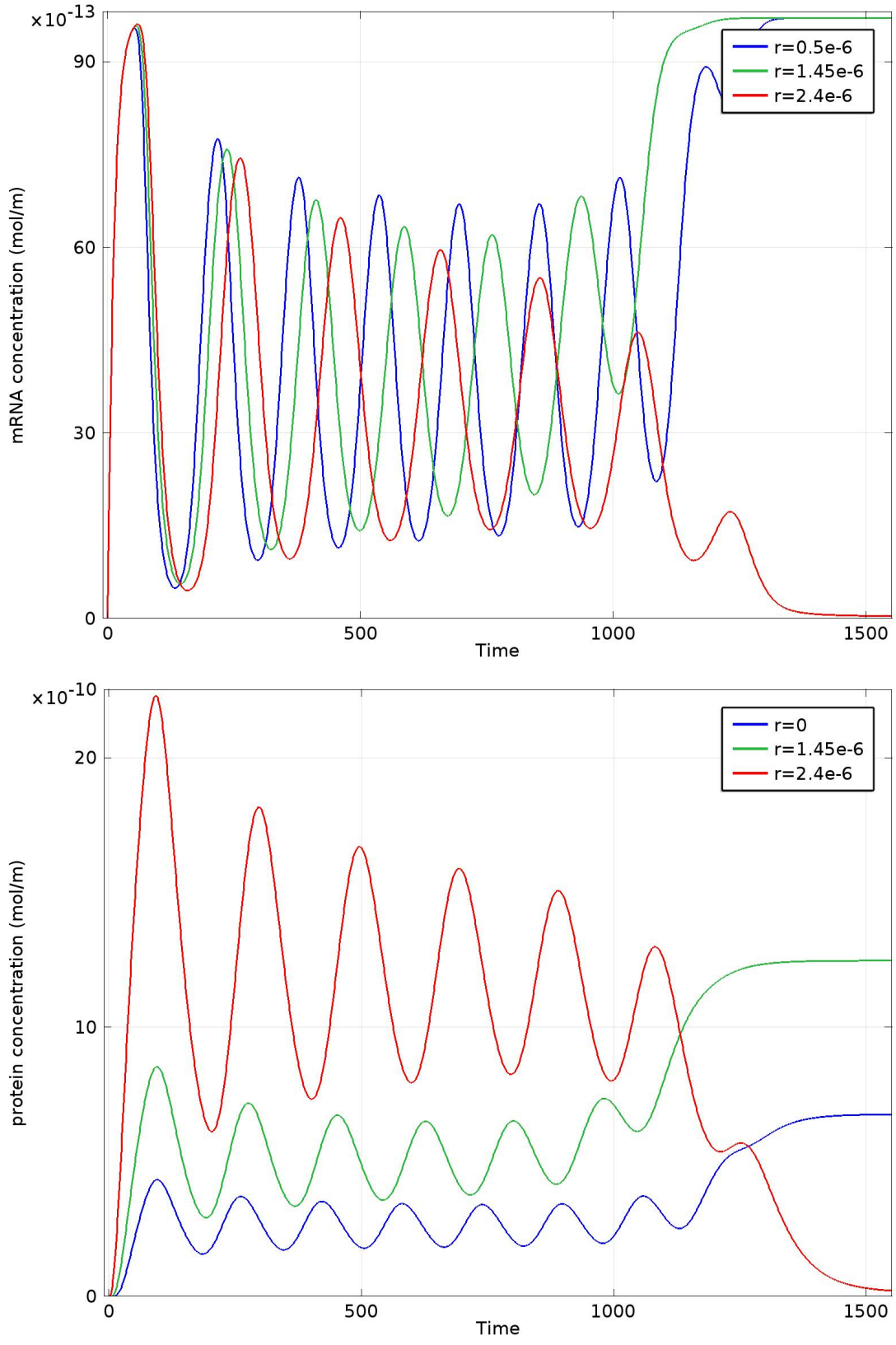

Figure 4.28: *Two-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Varying the equidistance of the two separate gene sites from the origin along the x-axis.* $D = 5 \times 10^{-14} g_1 = (r\mu m, 0\mu m), g_2 = (-r\mu m, 0\mu m)$*:- blue* $r = 0.5$*, Green* $r = 1.45$ *and red* $r = 2.4$

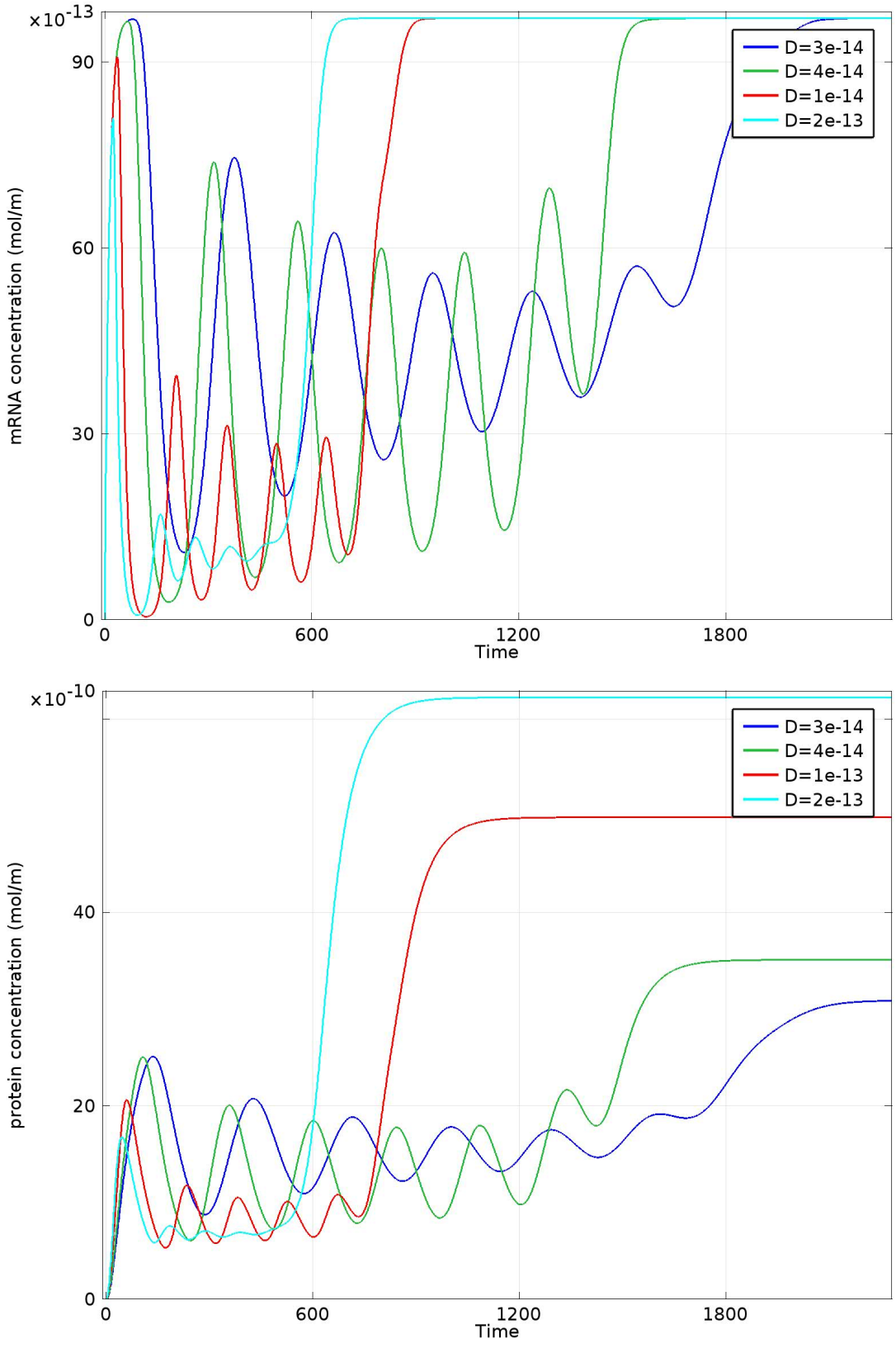

Figure 4.29: *Two-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Two separate gene sites at position close to the nuclear membrane, translated along the positive and negative x-axis: Varying D.* $g_1 = (2.4\mu m, 0\mu m), g_2 = (-2.4\mu m, 0\mu m)$*:- blue* $D = 3 \times 10^{-14}$*, Green* $D = 4 \times 10^{-14}$*, red* $D = 1 \times 10^{-14}$ *and cyan* $D = 2 \times 10^{-13}$

We now consider varying the position of the gene sites as a neighbouring cluster, such that they remain neighbours i.e. vary their position in the y direction only, such that the symmetry of the cell is broken down even further. The result of this is that the spatial-temporal dynamics is abruptly affected, particularly in length of the transient time, (see figure 4.30). As symmetry is further broken, the transient time of dynamics comes to an end at a faster rate. For the cluster of genes near the nuclear membrane $r = 2.4\mu m$, early phase oscillations were revived for a short time, for a small range of decreased diffusion coefficients, after which, steady states were observed, (see figure 4.31).

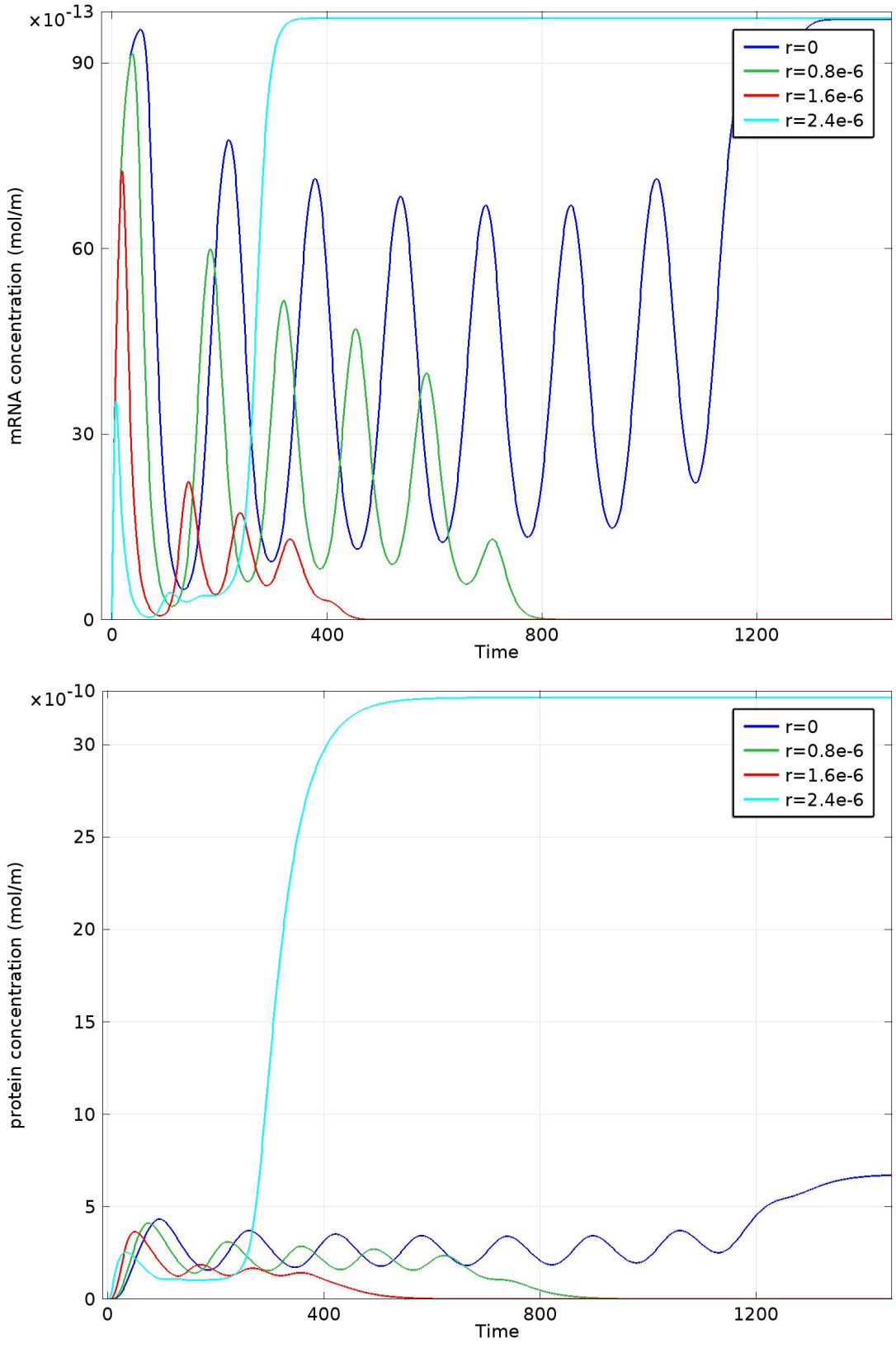

Figure 4.30: *TTwo-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Varying the distance of the two separate but neighbouring gene sites from the origin.* $g_1 = (0.5\mu m, r\mu m)$, $g_2 = (-0.5\mu m, r\mu m)$*:- blue* $r = 0$, *Green* $r = 0.8$, *red* $r = 1.6$ *and Cyan* $y = 2.4$

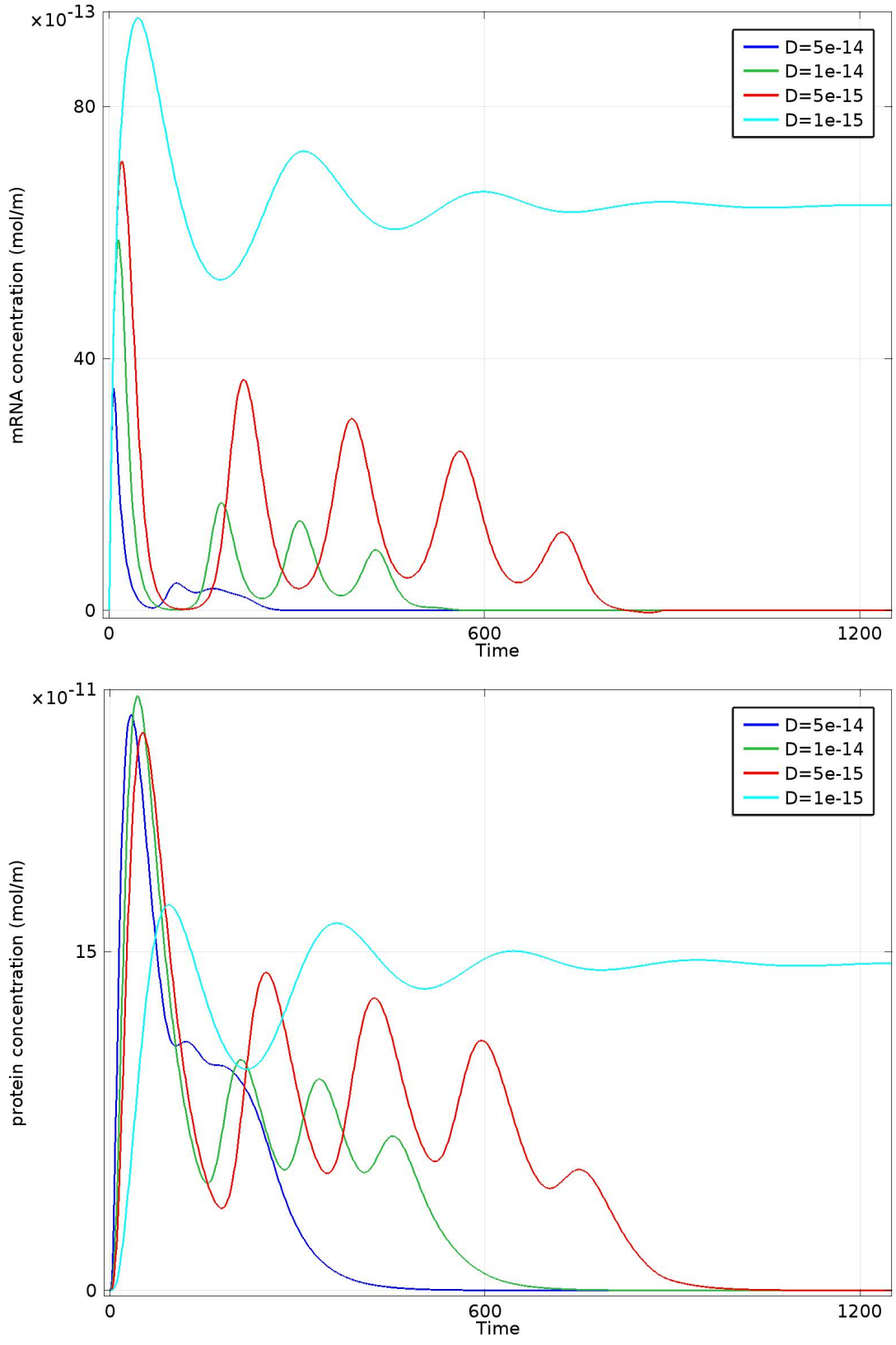

Figure 4.31: *Two-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Two separate but neighbouring gene sites positioned close the nuclear membrane.* $g_1 = (0.5\mu m, 2.4\mu m)$, $g_2 = (-0.5\mu m, 2.4\mu m)$*:-* *blue* $D = 5 \times 10^{-14}$, *Green* $D = 1 \times 10^{-14}$, *red* $D = 5 \times 10^{-15}$ *and Cyan* $D = 1 \times 10^{-15}$

4.4 The three gene repressilator

We now consider the three gene repressilator: $n = 3$ of the n-gene repressilator reaction diffusion equations.

4.4.1 Shared gene site

We consider the same geometry as the Hes1 system. The single gene site at the origin of the domain is where transcription of mRNA1, mRNA2 and mRNA3 can occur.

For the default value of $D = 5 \times 10^{-14}$ synchronous spatial-temporal oscillations were recorded for all mRNA and protein species for a transient time After this time, (here $t = 4000$), the phase change resulted in out of phase spatial-temporal oscillations, (see figures 4.32. This second phase is what is seen in the temporal dynamics of the Elowitz & Leibler (2000) ODE model. Due to the fact that the solution for the one-gene repressilator (with a shared gene site) is a solutions to this model, this suggests that the solution is unstable as it is expected that the initial in phase oscillations will last for all time, (Muller et al. 2006).

In consideration of varying the diffusion coefficient D we can see that there is a range of values, $5 \times 10^{-14} \leq D \leq 1 \times 10^{-13}$, to which sustained spatial-temporal oscillations occur, (see figures 4.33. Above and below this range oscillations die off. Within this oscillatory range, there are at most two phases of dynamics that can be observed. Oscillations that are in phase are observed first, and have a time span from two thousand to four thousand time units, depending on the specific D used. After this the dynamics switch to out of phase oscillations thereafter. Specifically, we see that as the diffusion coefficient is increased, the shorter the time it takes for the second phase of GRN dynamics to begin. Above and below this range of diffusion coefficients, oscillations are

damped quickly to a steady state solution.

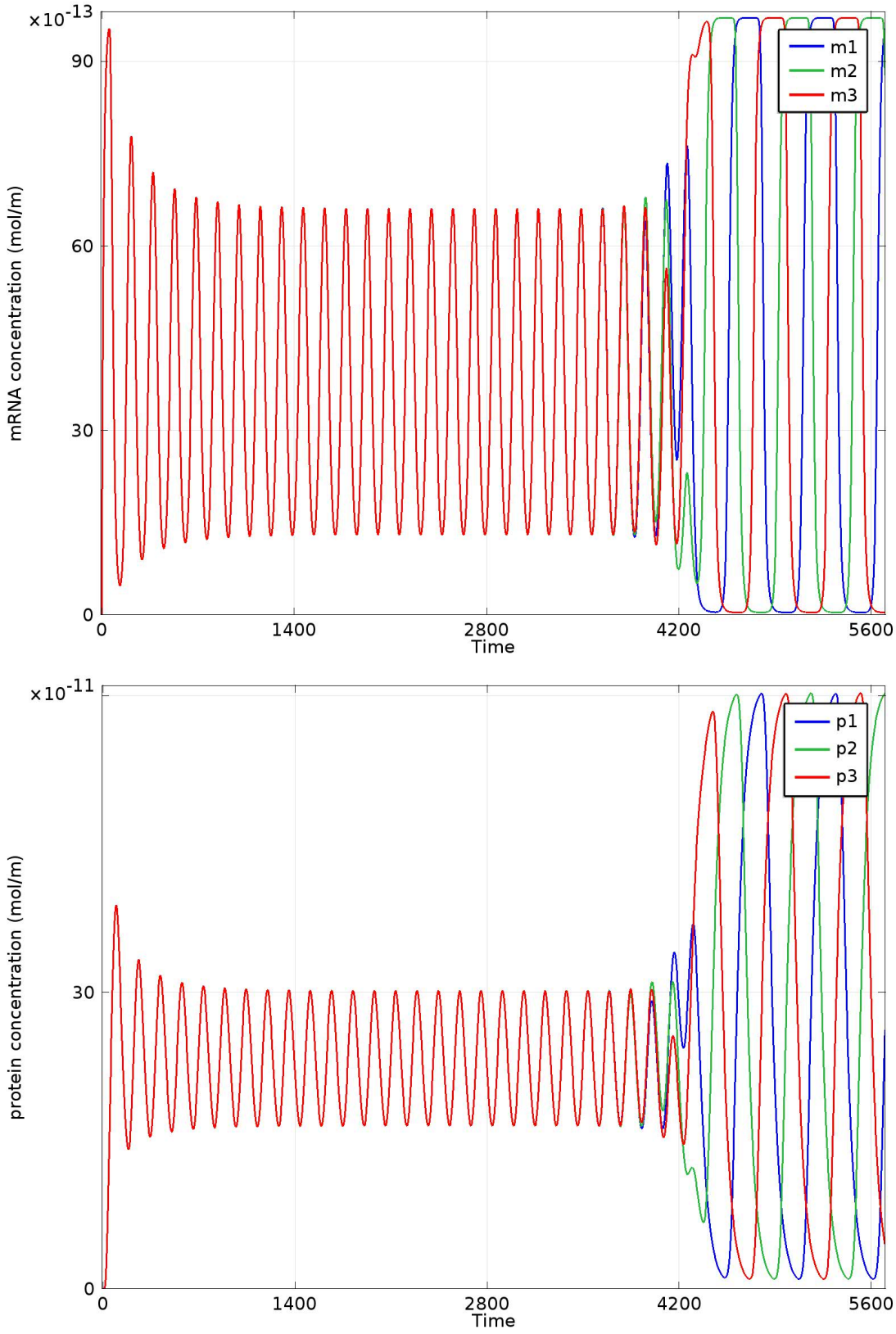

Figure 4.32: *Three-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Shared gene site at the origin,* $D = 5 \times 10^{-14}$

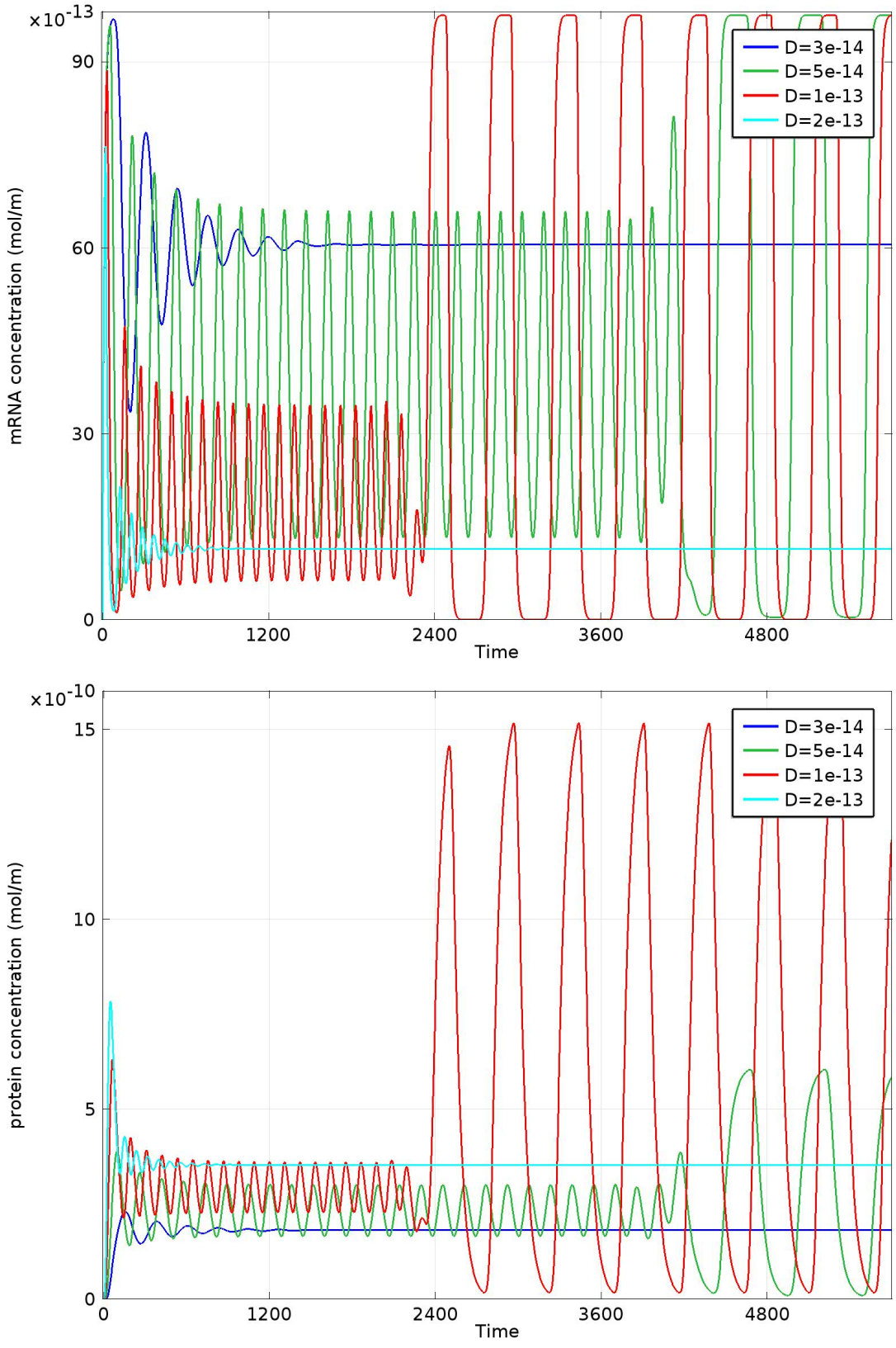

Figure 4.33: *Three-gene repressilator: Plots are of the integration of concentration levels of mRNA1 (top) and protein1 (bottom) over the entire cell against time. Varying D, shared gene site at the origin. green* $D = 2 \times 10^{-13}$, *Blue* $D = 1 \times 10^{-13}$, *red* $D = 5 \times 10^{-14}$ *and cyan* $D = 3 \times 10^{-14}$

In consideration of moving the location of the shared gene site, similar spatial-temporal dynamics are recorded (for the default value of $D = 5 \times 10^{-14}$), however, the time of the initial transient dynamics decreases from 4000 for $r = 0$ to 2500 for $r = 0.75$, (see figure 4.34. As a threshold value of $r = 1$ is reached, spatial-temporal oscillations, dampen and die. As symmetry is highly broken in the x and y axis here, the spatial-temporal dynamics is highly affected for small changes of the gene site location.

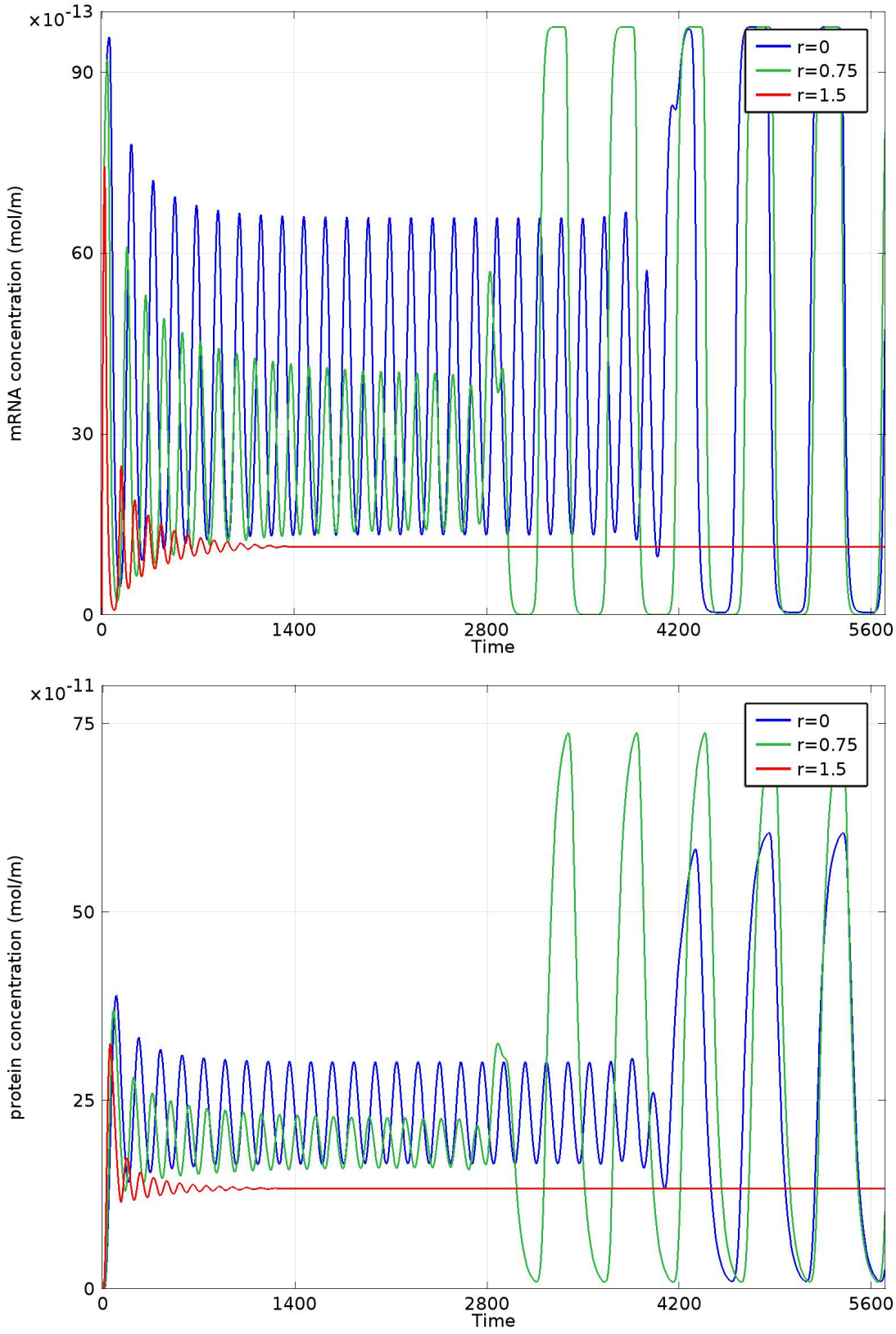

Figure 4.34: *Three-gene repressilator: Plots are of the integration of concentration levels of mRNA1 (top) and protein1 (bottom) over the entire cell against time. Varying the location of the shared gene site and keeping D fixed,* $D = 5 \times 10^{-14}$. *green* $r = 0\mu m$, *Blue* $r = 0.75\mu m$, *and red* $r = 1.5\mu m$

4.4.2 Individual gene sites

Let us now create individual gene sites within the nucleus. We will investigate the impact of clustered gene spatial location and a more spread out gene spatial location of the three genes. This will then lead us to make inferences about how the symmetry of the geometry affects the GRN dynamics. Each gene site is circular with radius 0.5μm.

We now look at the three genes clustered together at the origin. Consequently their spatial locations are then: $(x_1, y_1) = (0, \frac{1}{\sqrt{3}}\mu m)$, $(x_2, y_2) = (0.5\mu m, -\frac{1}{2\sqrt{3}}\mu m)$ and $(x_3, y_3) = (-0.5\mu m, -\frac{1}{2\sqrt{3}}\mu m)$ for gene 1, gene 2 and gene 3 respectively. Thus, we are choosing gene sites to be a distance of $1\mu m$ from each other, in a triangular configuration.

The same two phases of spatial-temporal oscillations are again seen for $D = 5 \times 10^{-14}$, (see figure 4.35). The early phase of spatial-temporal oscillations are all in phase for all three species of mRNA and protein. The later phase of spatial-temporal oscillations are seen to be out of phase. This initial transient phase has also been studied by Bennet et al. (2007). The difference being that the transient time comes to an end a lot earlier than with the single shared gene site at $t = 750$. The range of diffusion coefficients that achieve spatial-temporal oscillations are $4 \times 10^{-14} \leq D \leq 5 \times 10^{-13}$. As D is increased from within this range the transient time ends earlier and the amplitude of protein concentration levels is significantly increased. Within this range of D, mRNA amplitude levels remain consistent, (see figure 4.36). However, we suspect that for a high value of D the PDE system should reduce to the ODE system, such that oscillations will still be seen. Running the model again, with inequality in the initial conditions for the protein produced asynchronous oscillations, with the appearance of only a short kinky transient. Translating one of the gene sites out of the symmetrical configuration also had the effect of omitting the initial transient phase of in-phase oscillations.

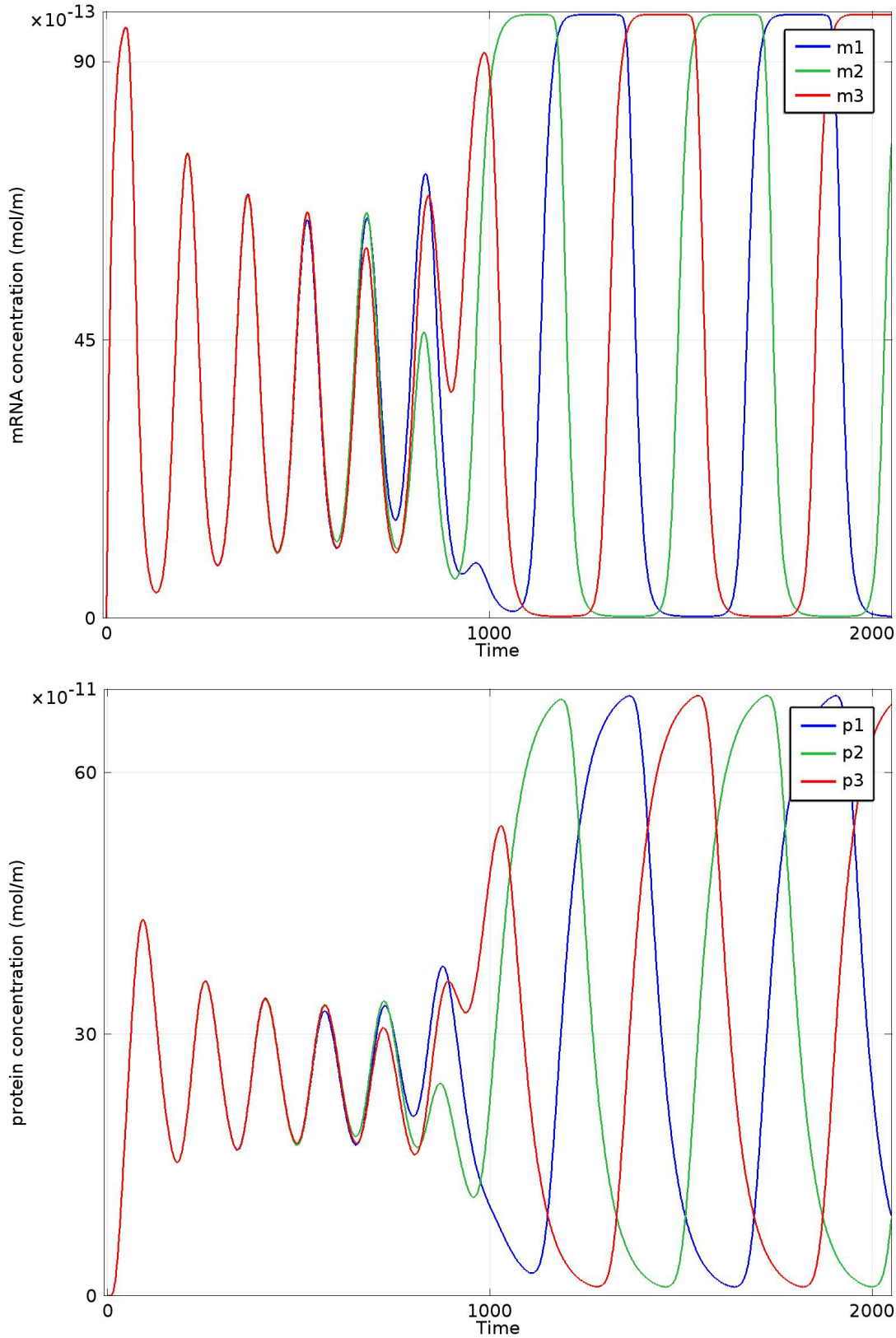

Figure 4.35: *Three-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Individual gene sites clustered together at the origin,* $D = 5 \times 10^{-14}$

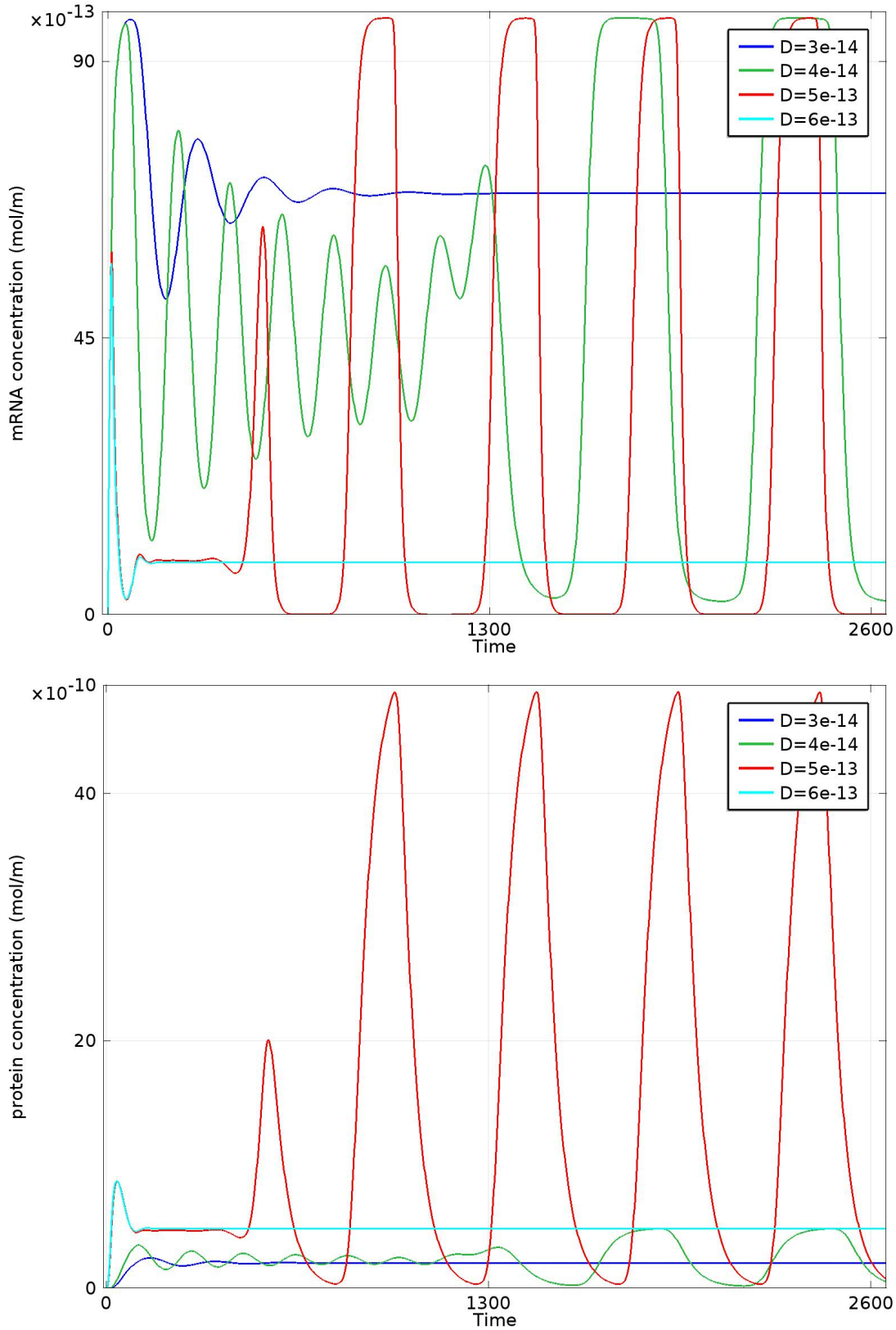

Figure 4.36: *Three-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Individual gene sites, clustered together at the origin: Varying D*

In consideration of varying the equidistance of the gene sites from their original separation of $1\mu m$ to $2.4\mu m$, no further symmetry is broken.

Thus, we let $(x_1, y_1) = (0\mu m, 2.4\mu m)$, $(x_2, y_2) = (2.4\cos(\pi/6)\mu m, -2.4\sin(\pi/6)\mu m)$ and $(x_3, y_3) = (-2.4\cos(\pi/6)\mu m, -2.4\sin(\pi/6)\mu m)$ be the coordinates of the centres of gene 1, gene 2 and gene 3 respectively. Below is a temporal plot of total concentration of the mRNA and protein species for the model with gene sites equidistant away from each other, close to the nuclear membrane. The spatial-temporal dynamics observed did not change from that seen of the separate gene sites clustered together at the origin. The spatial-temporal dynamics differ from that of the single gene variation of r as here, oscillations were not lost from the greater separation of gene location. The early phase of synchronous oscillations and the later phase of asynchronous oscillations, remains relatively unchanged. This may be partly due to the preserved symmetry. However, in consideration of varying D, the range of spatial-temporal oscillations recorded was for a range shifted down to a slightly smaller value i.e. $3 \times 10^{-14} \leq D \leq 3 \times 10^{-13}$.

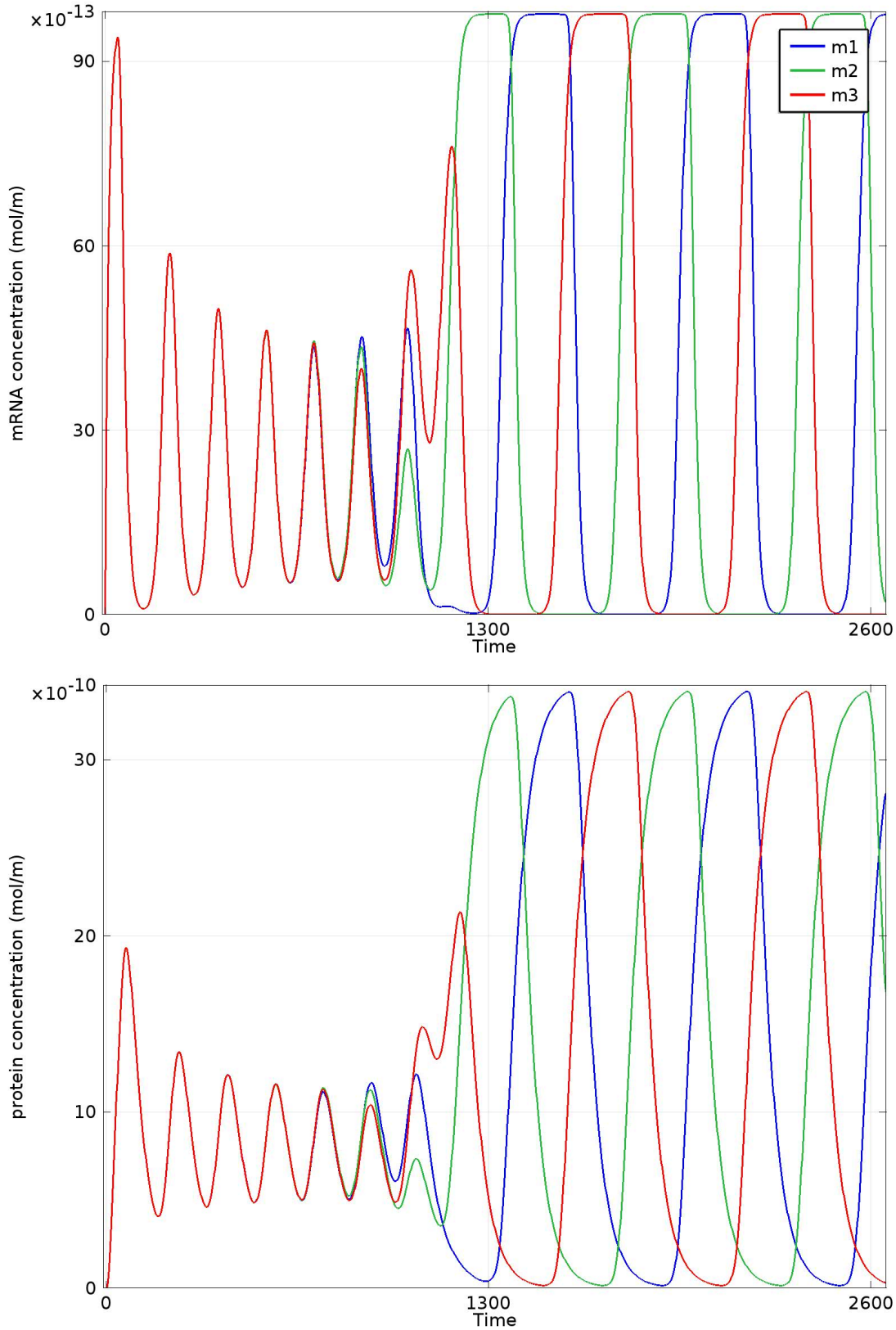

Figure 4.37: *Three-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Individual gene sites, each close to the nuclear membrane but translated along the positive and negative x-axis,* $D = 5 \times 10^{-14}$

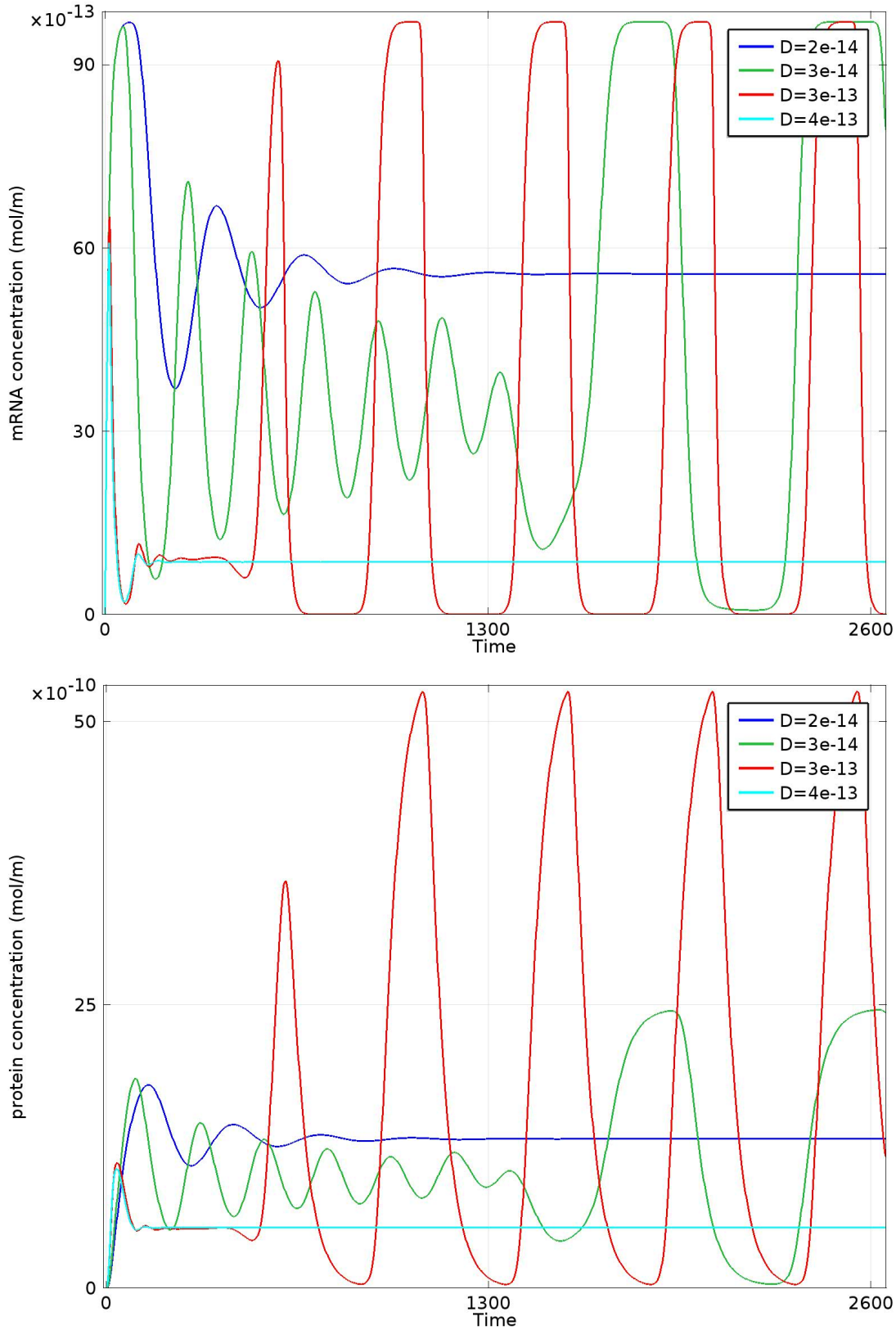

Figure 4.38: *The three gene repressilator: Individual gene sites, each at the nuclear membrane, equidistant from each other and the origin: Varying D*

We now consider varying the distance of the gene sites as a cluster from the origin, such that their locations are now given by $(x_1, y_1) = (0, \frac{1}{\sqrt{3}} + y\mu m)$, $(x_2, y_2) = (0.5\mu m, -\frac{1}{2\sqrt{3}} + y\mu m)$ and $(x_3, y_3) = (-0.5\mu m, -\frac{1}{2\sqrt{3}} + y\mu m)$ for gene 1, gene 2 and gene 3 respectively.

This change in gene location breaks symmetry further in both the x and y axis. The larger y becomes, the smaller the transient time of synchronous spatial-temporal oscillations and the larger the amplitude of protein concentration levels becomes. This differs from the single gene variation in r as oscillations are completely lost as r is increased but for individual gene sites, they are not. However, the difference in the gene sites being separated but equidistantly and the shift of the cluster in the y direction is that the transient time of synchronous oscillations decreases very quickly and the second phase of asynchronous oscillations dominates. The only first phase to be seen is a single first peak (although at different concentration levels) in the mRNA and in the protein concentrations. This can be explained by the fact that gene one is now closer to the cytoplasm than gene sites two and three. Thus, creating a difference in their expression levels right from the beginning.

A further observation we make about the three gene repressilator, when the symmetry of the geometry is skewed i.e the gene sites are clustered together at the nuclear membrane is that we do not seem to have an upper diffusion coefficient limit. This is as expected as we expect the PDE model to imitate the ODE model for a high value of D. It appears that our model is more stable when there is an asymmetry in the system. Whether that be in the different distances between different gene sites and the cytoplasm, differing initial conditions and parameter values.

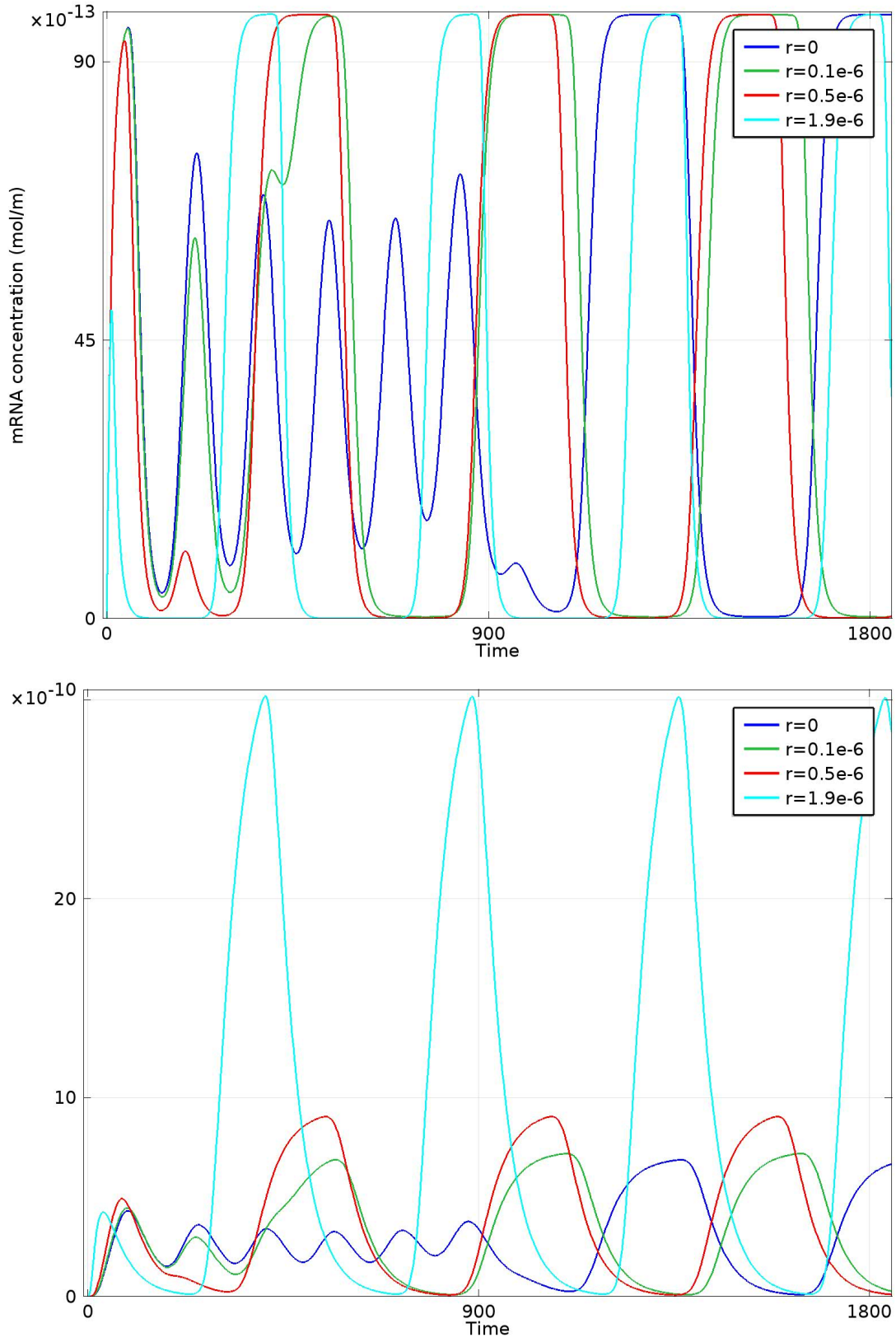

Figure 4.39: *Three-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Individual gene sites clustered together, varying their distance from the origin,* $D = 5 \times 10^{-14}$

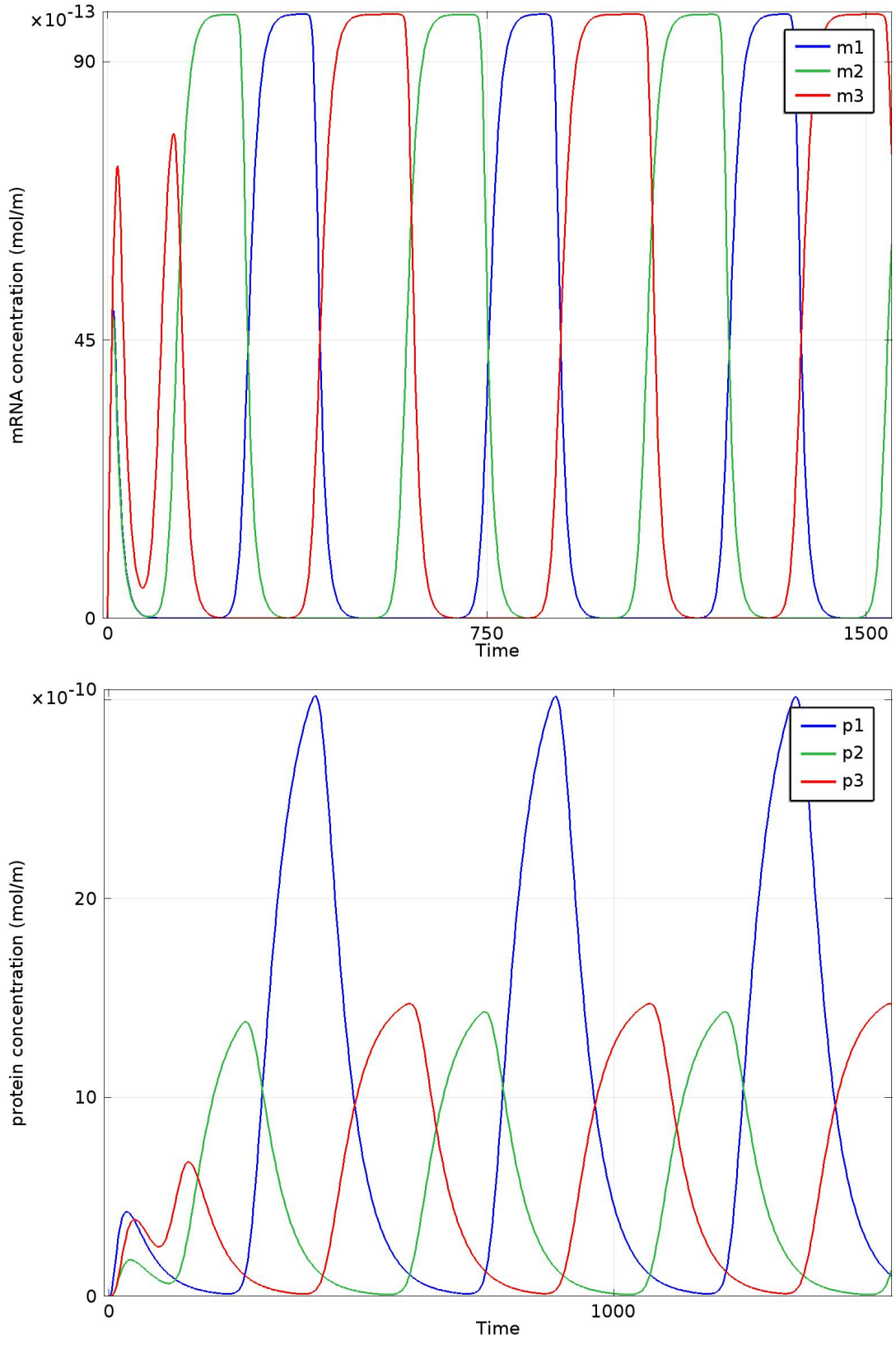

Figure 4.40: *Three-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Individual gene sites clustered together and close to the nuclear membrane,* $D = 5 \times 10^{-14}$

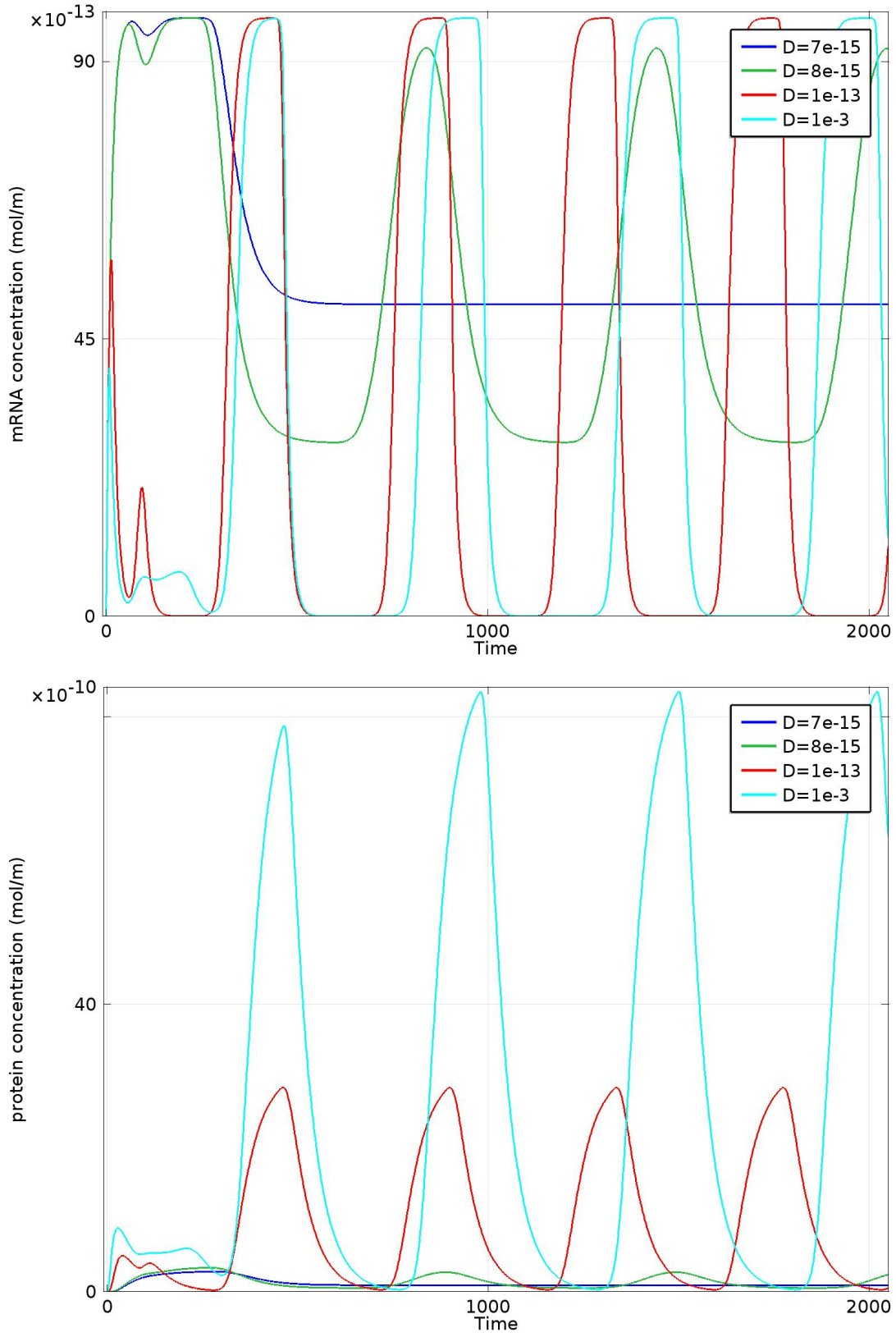

Figure 4.41: *Three-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Individual gene sites clustered together and close to the nuclear membrane: Varying D*

4.5 The four gene repressilator

We now consider the four gene repressilator: $n = 4$ of the n-gene repressilator reaction diffusion equations.

4.5.1 A single shared gene site

We find that the results of the four gene repressilator, with a single shared gene site, are similar to that of the two gene repressilator, with a single shared gene site. The early and later phase dynamics, transient time and change in concentration amplitude are all relatively the same. This still holds after consideration of varying the gene site location and in varying the diffusion coefficient. Hence, under the same parameter values as before and for a similar set of diffusion and gene site location parameter ranges, we find that the first phase of dynamics consists of synchronous spatial-temporal oscillations and the second phase to consist of steady states solutions, (see figures 4.42, 4.45, 4.46 and 4.47. These steady state solutions are again in agreement with long term steady state solution of ODE models, (Muller et al. 2006).

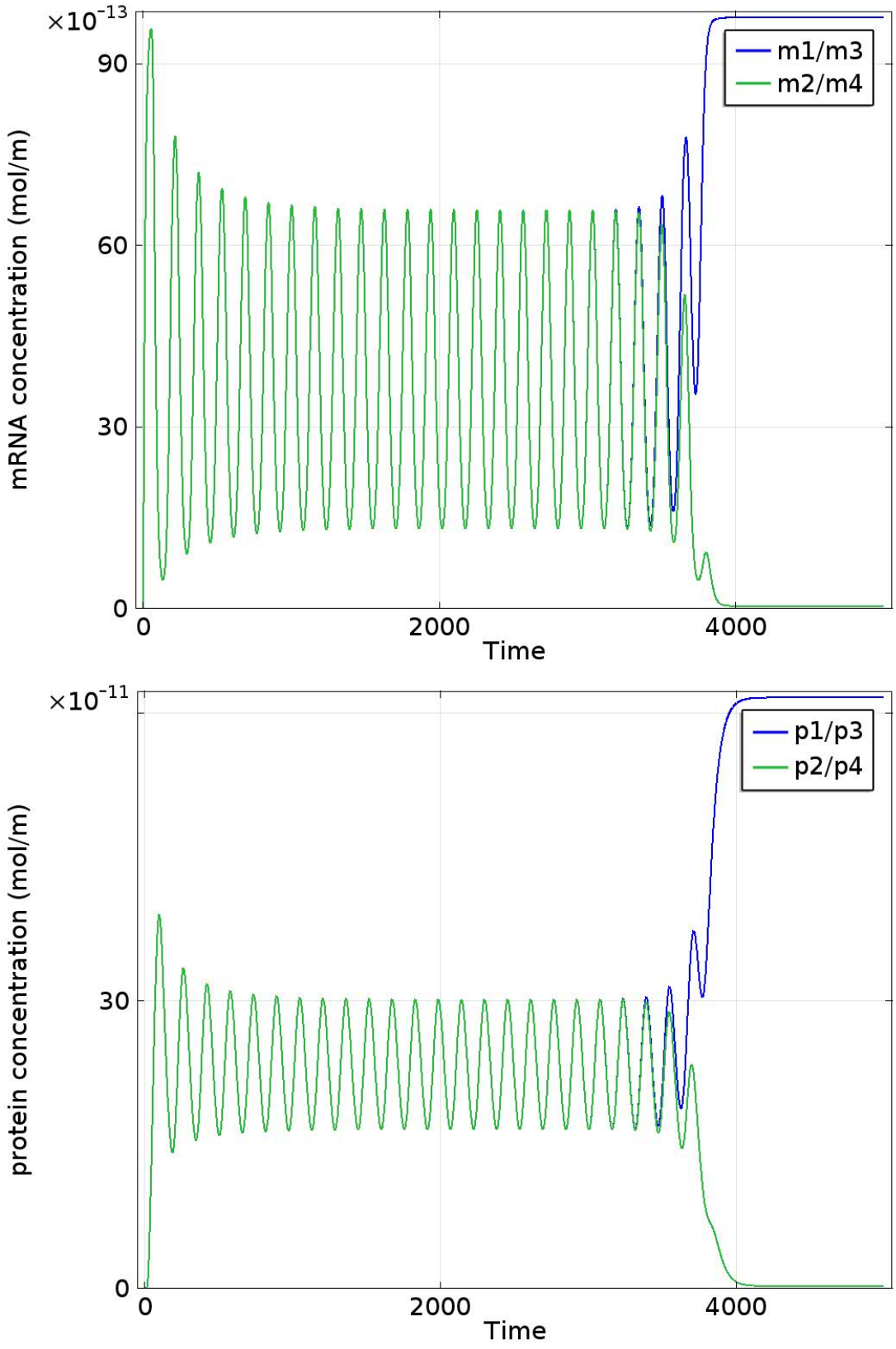

Figure 4.42: *Four-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. A single central gene site and* $D = 5 \times 10^{-14}$

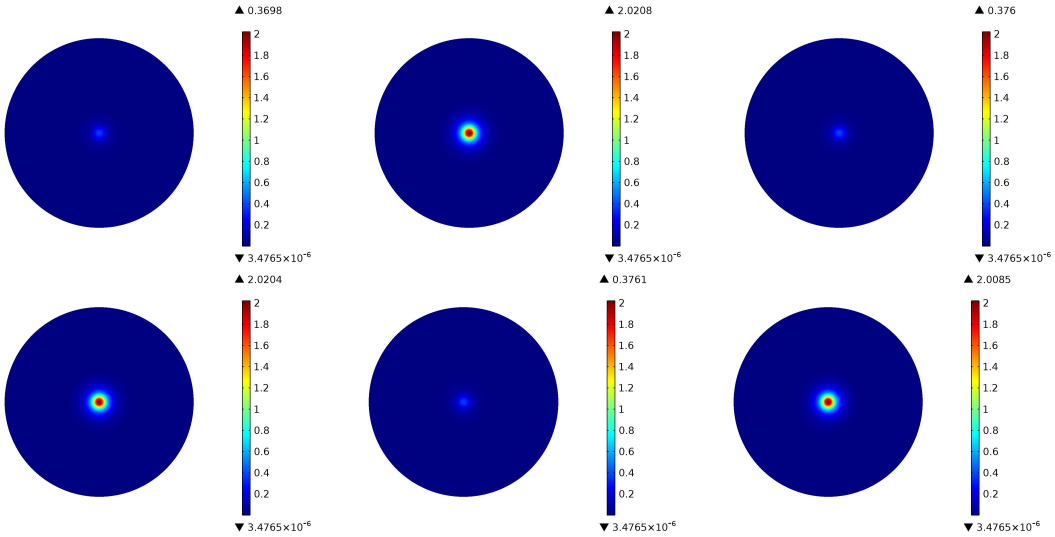

Figure 4.43: *The four gene repressilator with a single central gene site. Spatial snapshots of mRNA1.* $D = 5 \times 10^{-14}$

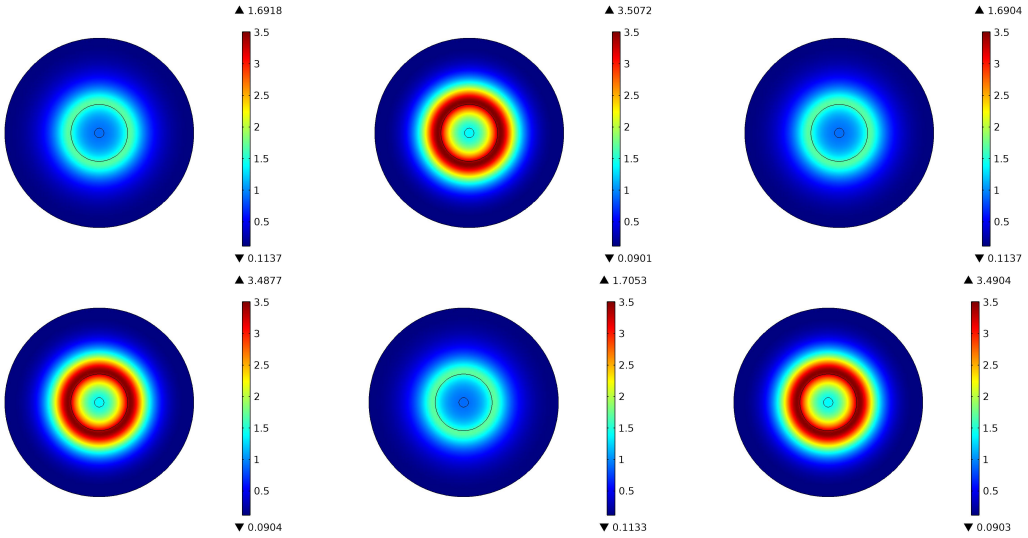

Figure 4.44: *The four gene repressilator with a single central gene site. Spatial snapshots of protein1.* $D = 5 \times 10^{-14}$

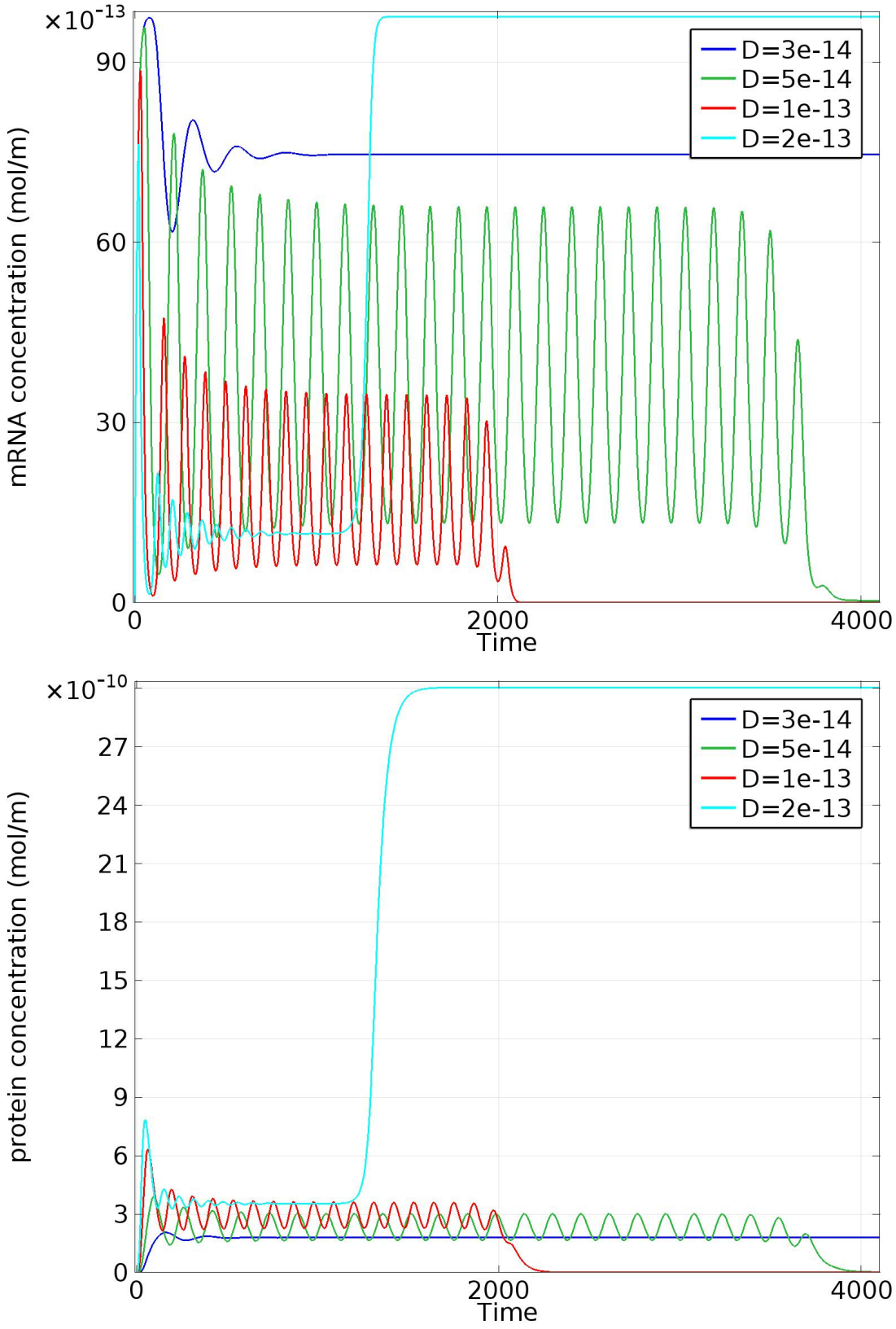

Figure 4.45: *Four-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. One shared gene site at the origin. Varying the diffusion coefficient. Both graphs are time series of total concentration of species. Top: mRNA, Bottom: protein.*

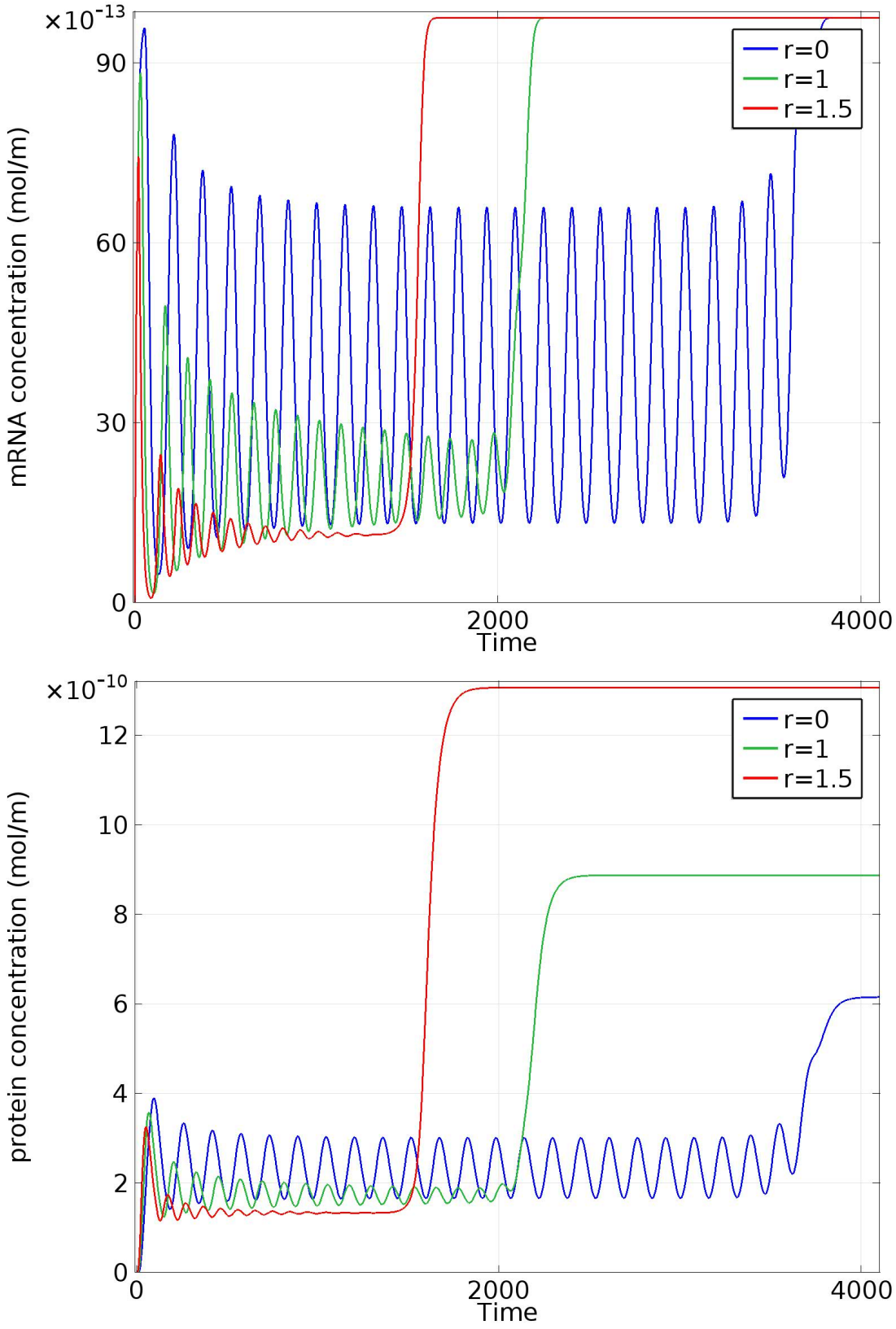

Figure 4.46: *Four-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Varying the location of the single gene site. Both graphs are time series of total concentration of species. Top: mRNA, Bottom: Protein.* $D = 5 \times 10^{-14}$.

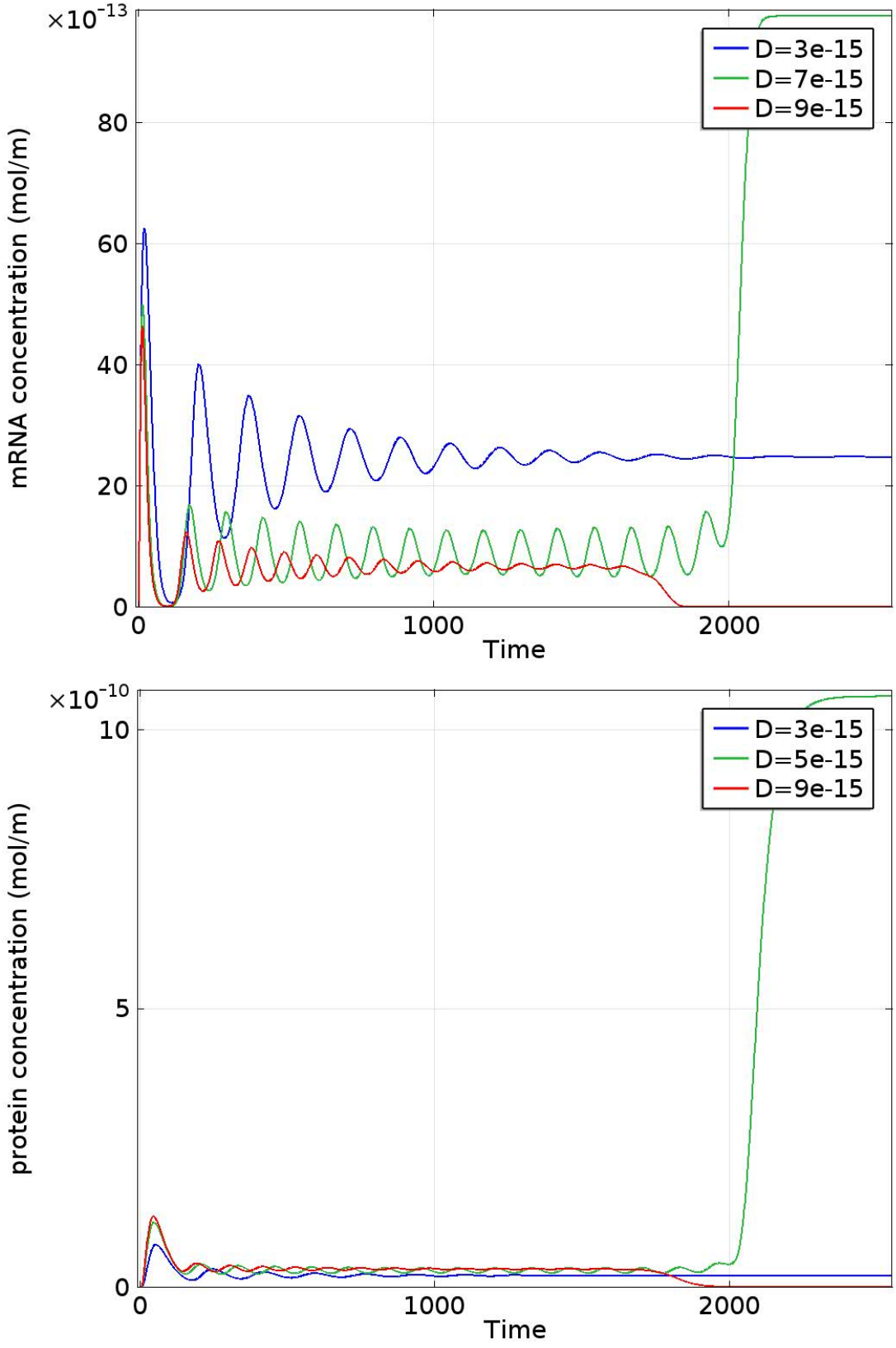

Figure 4.47: *Four-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. A single central gene site close to the nuclear membrane, $r = 2.4\mu m$.*

4.5.2 Four separate individual gene sites

We now consider four individual gene sites existing as a non intersecting cluster at the centre of the nucleus. The location of gene 1, gene 2, gene 3 and gene 4 being, $(x_1, y_1) = (0\mu m, \frac{1}{\sqrt{2}}\mu m)$, $(x_2, y_2) = (\frac{1}{\sqrt{2}}\mu m, 0\mu m)$, $(x_3, y_3) = (0\mu m, -\frac{1}{\sqrt{2}}\mu m)$ and $(x_4, y_4) = (-\frac{1}{\sqrt{2}}\mu m, 0\mu m)$ respectively. Thus, they are in a square configuration of sides of length $1\mu m$.

We find that the results of the four gene repressilator, with individual gene sites, are similar to that of the two gene repressilator, with individual gene sites. This might be due to the possibility that our model might not be independent of mesh size or it could be in further support of the results by Strelkowa & Barahona (2010). The transient time of oscillatory dynamics and change in concentration amplitude are all relatively the same. This still holds after consideration of varying the gene site location and in varying the diffusion coefficient. Hence, under the same parameter values as before and for a similar set of diffusion and gene site location parameter ranges, we find that the first phase of dynamics consist of synchronous spatial-temporal oscillations and the second phase to consist of steady states solutions.The change captured between the single shared gene site model to the individual gene sites model also reflects that of the change seen from a single shared gene to individual gene sites in the two gene repressilator model. Thus, the first phase of spatial-temporal oscillations for single shared gene lasts considerably longer than for the four individual origin clustered genes model. Furthermore, the initial transient time decreased further for an asymmetrical gene site cluster at the nuclear membrane. Also, again for changes to the diffusion coefficient, we see that the transient time, period of oscillation, amplitude of concentration, particularly of protein is deeply altered. However, in the case of clustered gene sites at the nuclear membrane, the four gene repressilator with individual gene sites model's first phase of (damped) oscillatory dynamics have a much larger protein amplitude.

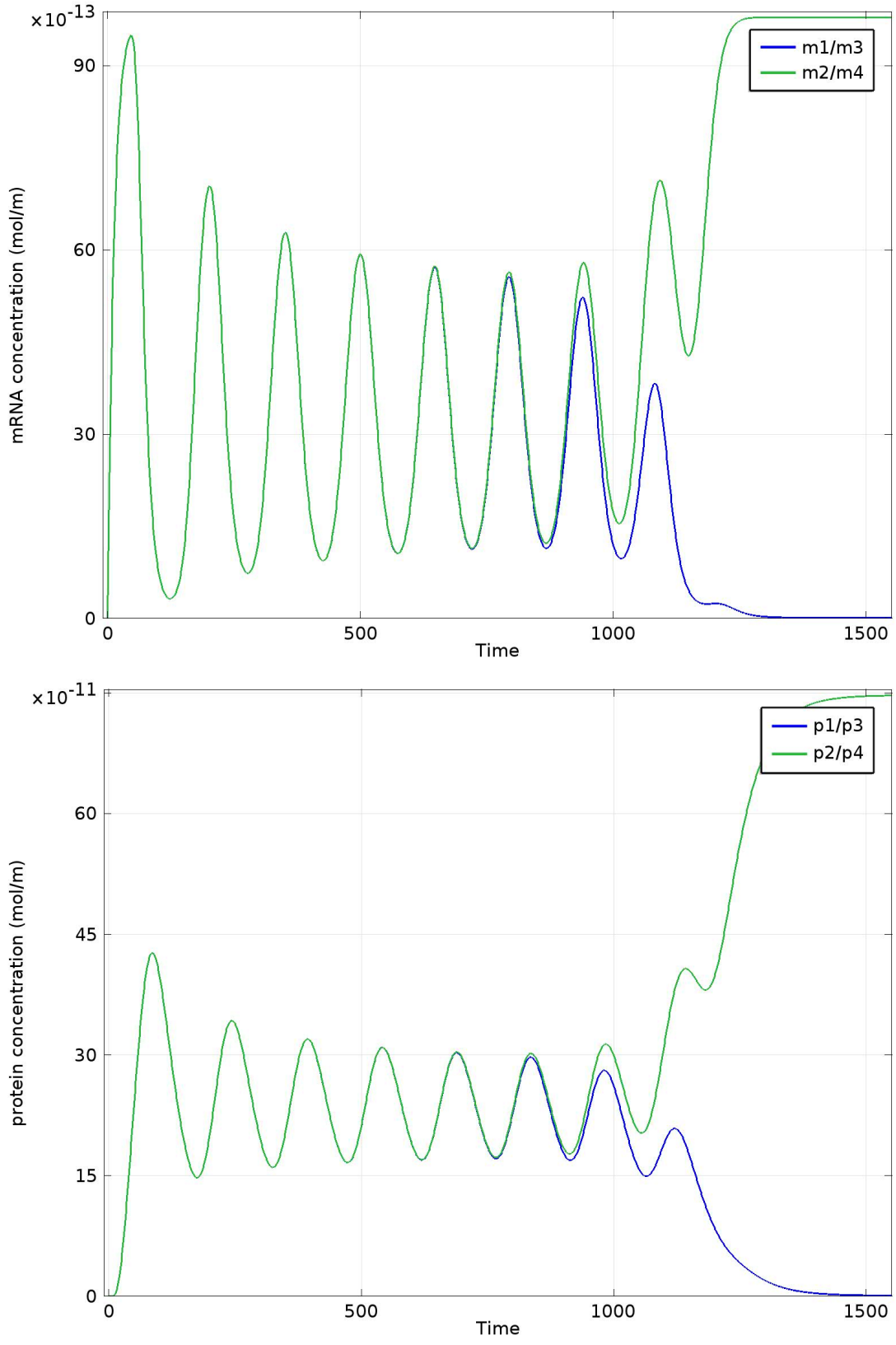

Figure 4.48: *Four-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Individual gene sites, clustered together at the origin, with specific gene site location being,* $g_{1/3} = (+/-\frac{1}{\sqrt{2}}, 0)$ *and* $g_{2/4} = (0, +/-\frac{1}{\sqrt{2}})$*. Both graphs are time series of total concentration of species.* $D = 5 \times 10^{-14}$*. Top: mRNA, Bottom: Protein.*

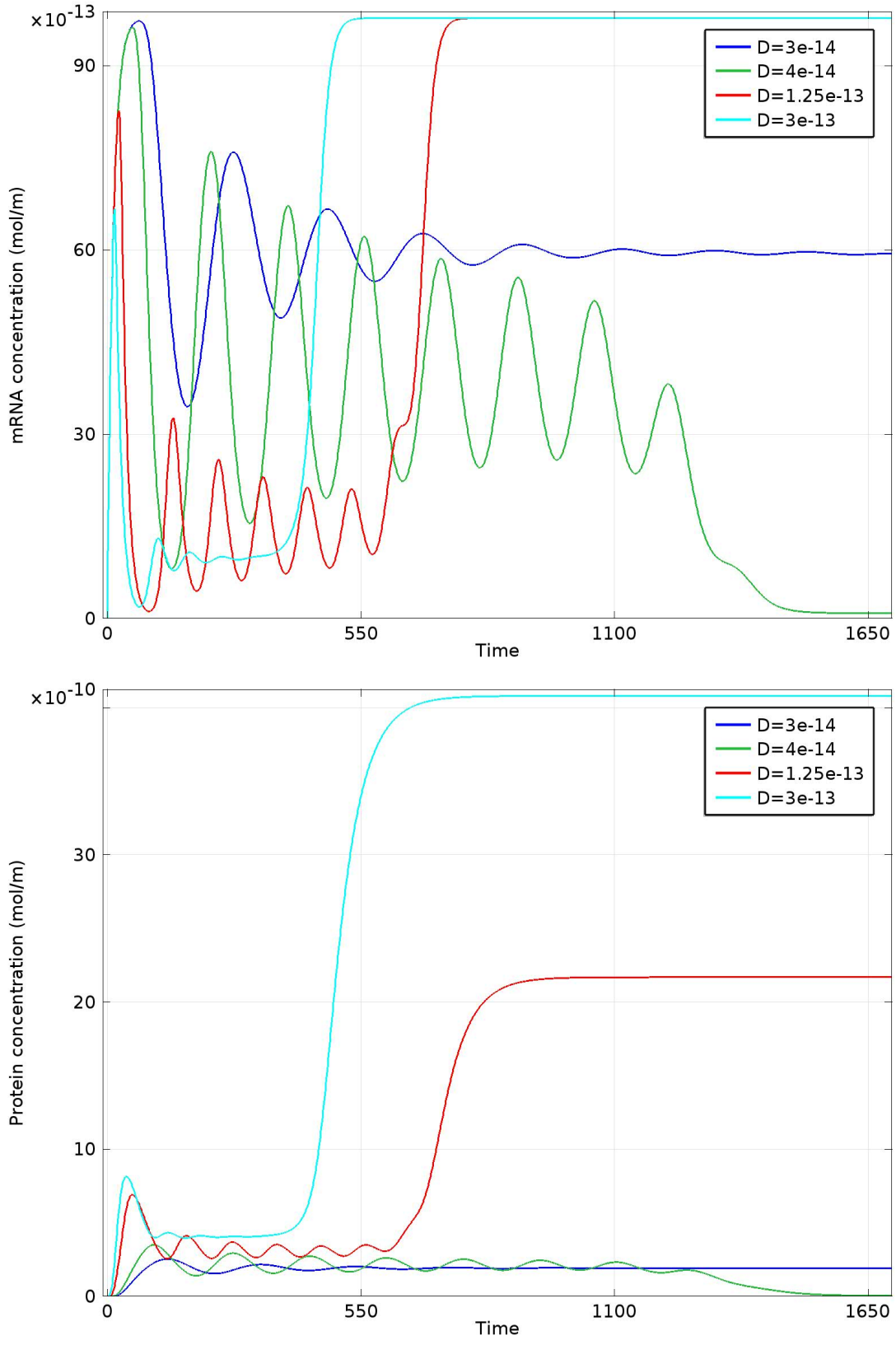

Figure 4.49: *Four-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Individual gene sites, clustered together at the origin, with specific gene site location being,* $g_{1/3} = (+/-\frac{1}{\sqrt{2}}, 0)$ *and* $g_{2/4} = (0, +/-\frac{1}{\sqrt{2}})$*: Varying D.*

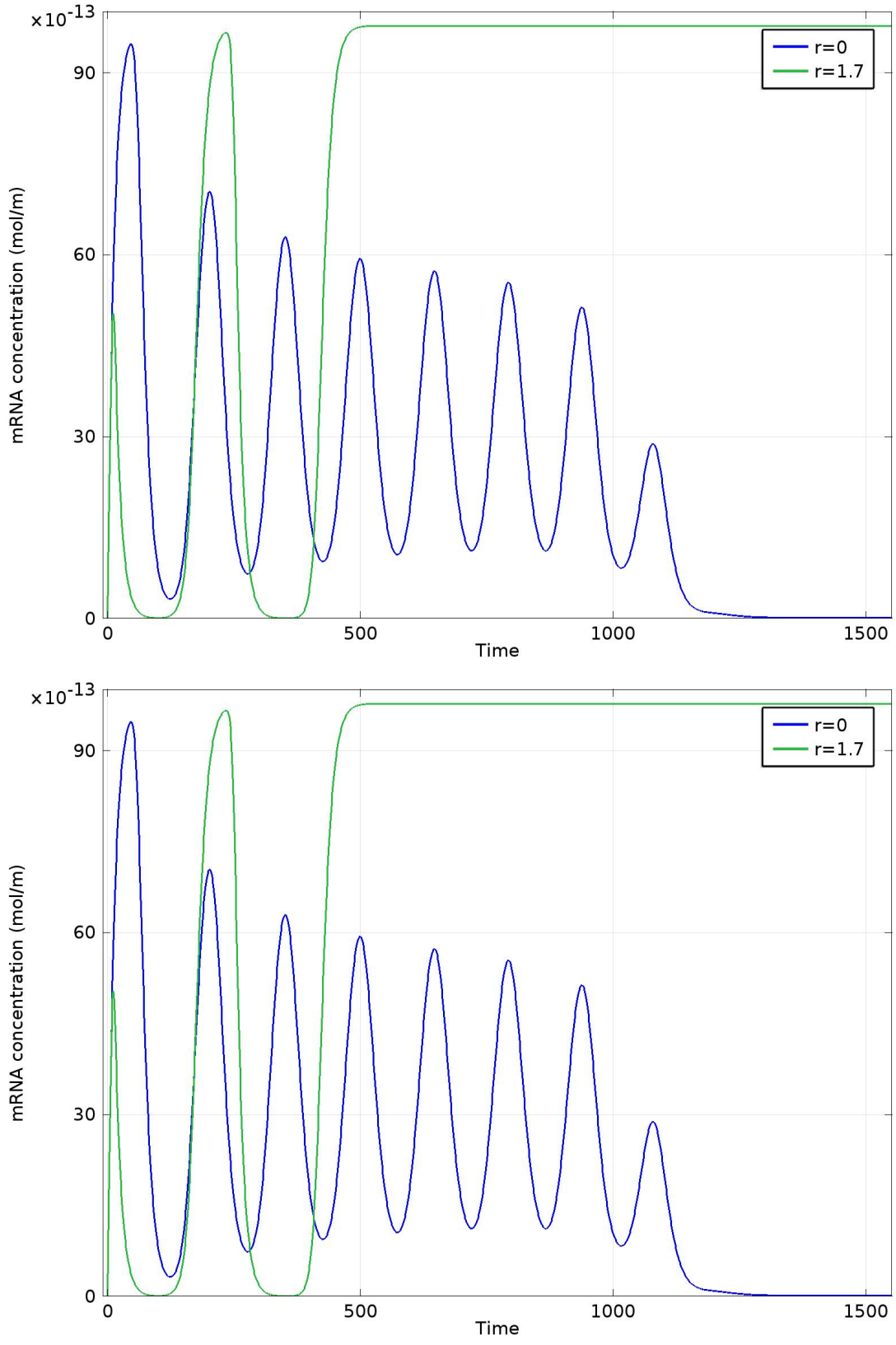

Figure 4.50: *Four-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Varying the position of cluster of individual gene sites, with specific gene site location being,* $g_{1/3} = (+/-\frac{1}{\sqrt{2}}, r)$ *and* $g_{2/4} = (0, r+/-\frac{1}{\sqrt{2}})$.

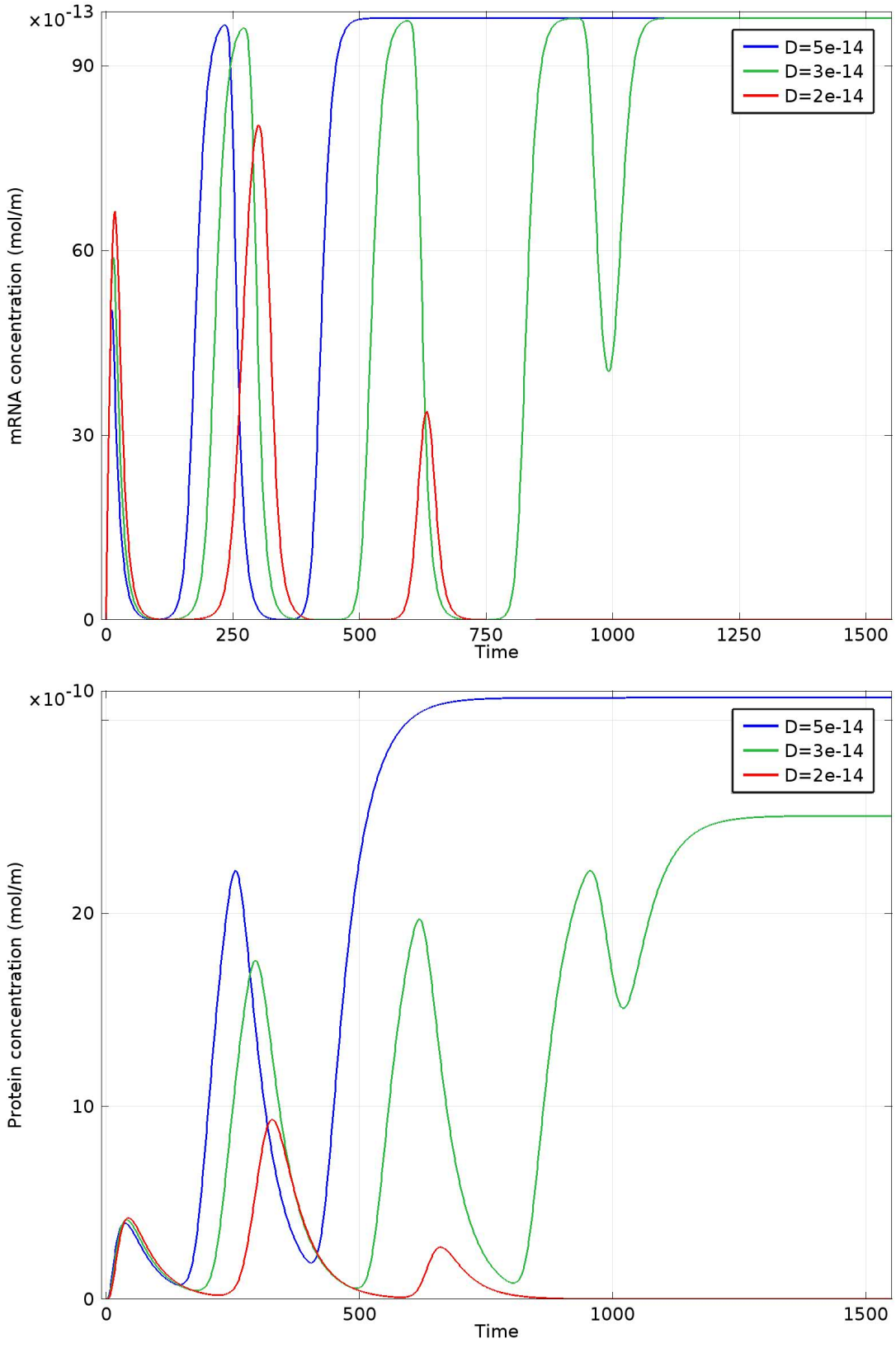

Figure 4.51: *Four-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Individual gene sites, clustered together near the nuclear membrane, with specific gene site location being,* $g_{1/3} = (+/- \frac{1}{\sqrt{2}}\mu m, 1.7\mu m)$ *and* $g_{2/4} = (0\mu m, (1.7 +/- \frac{1}{\sqrt{2}})\mu m)$*: Varying D.*

4.6 The five gene repressilator

We now consider the five gene repressilator: $n = 5$ of the n-gene repressilator reaction diffusion equations.

4.6.1 A single shared gene site

We find that the results of the five gene repressilator, with a single shared gene site, are similar to that of the three gene repressilator, with a single shared gene site. The early and later phase dynamics, transient time and change in concentration amplitude are all relatively the same. This still holds after consideration of varying the gene site location and in varying the diffusion coefficient. Hence, under the same parameter values as before and for a similar set of diffusion and gene site location parameter ranges, we find that the transient time of initial dynamics consist of synchronised spatial-temporal oscillations and the second phase to consist of asynchronous spatial-temporal oscillations.

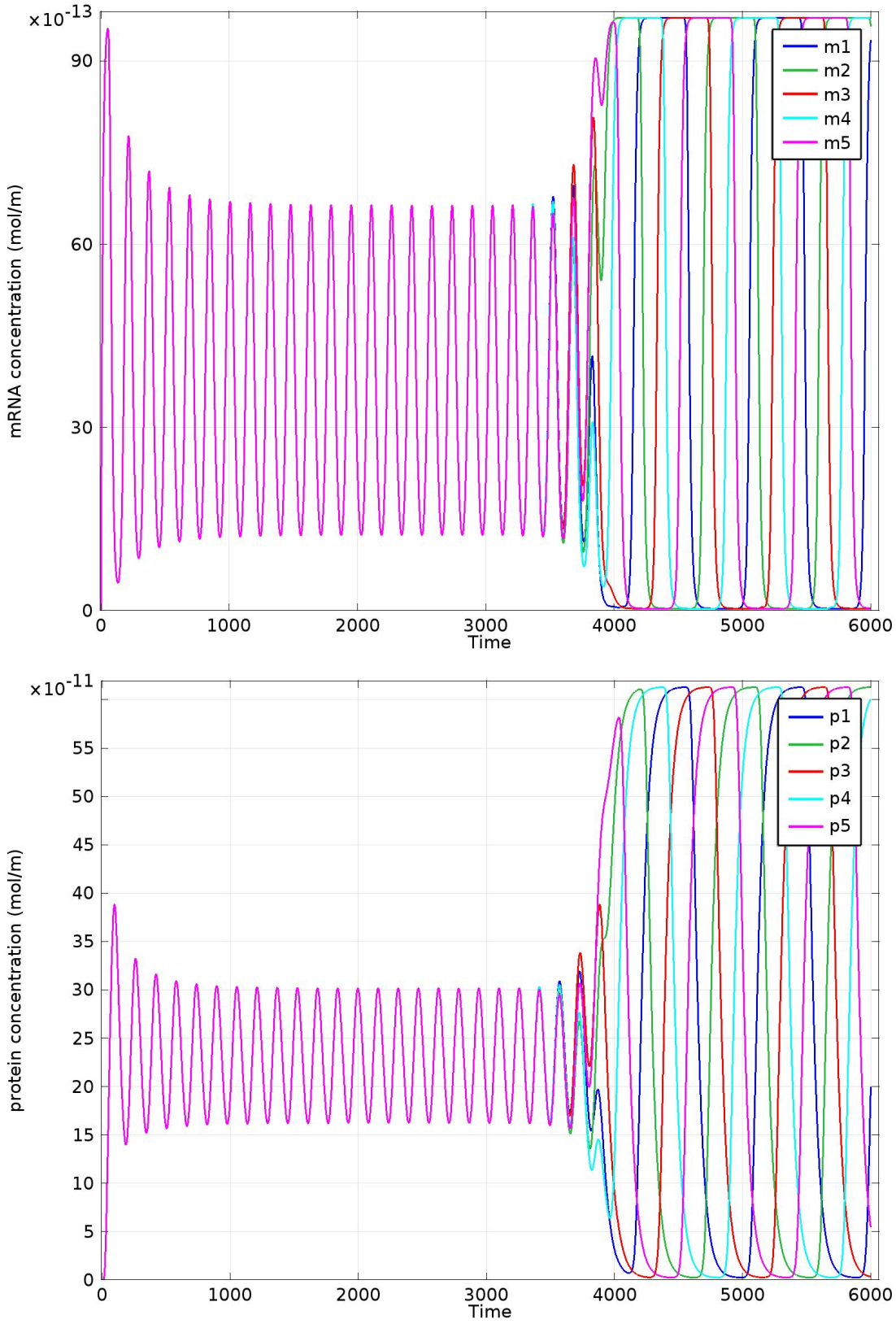

Figure 4.52: *Five-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. A shared gene site at the origin:* $D = 5 \times 10^{-14}$

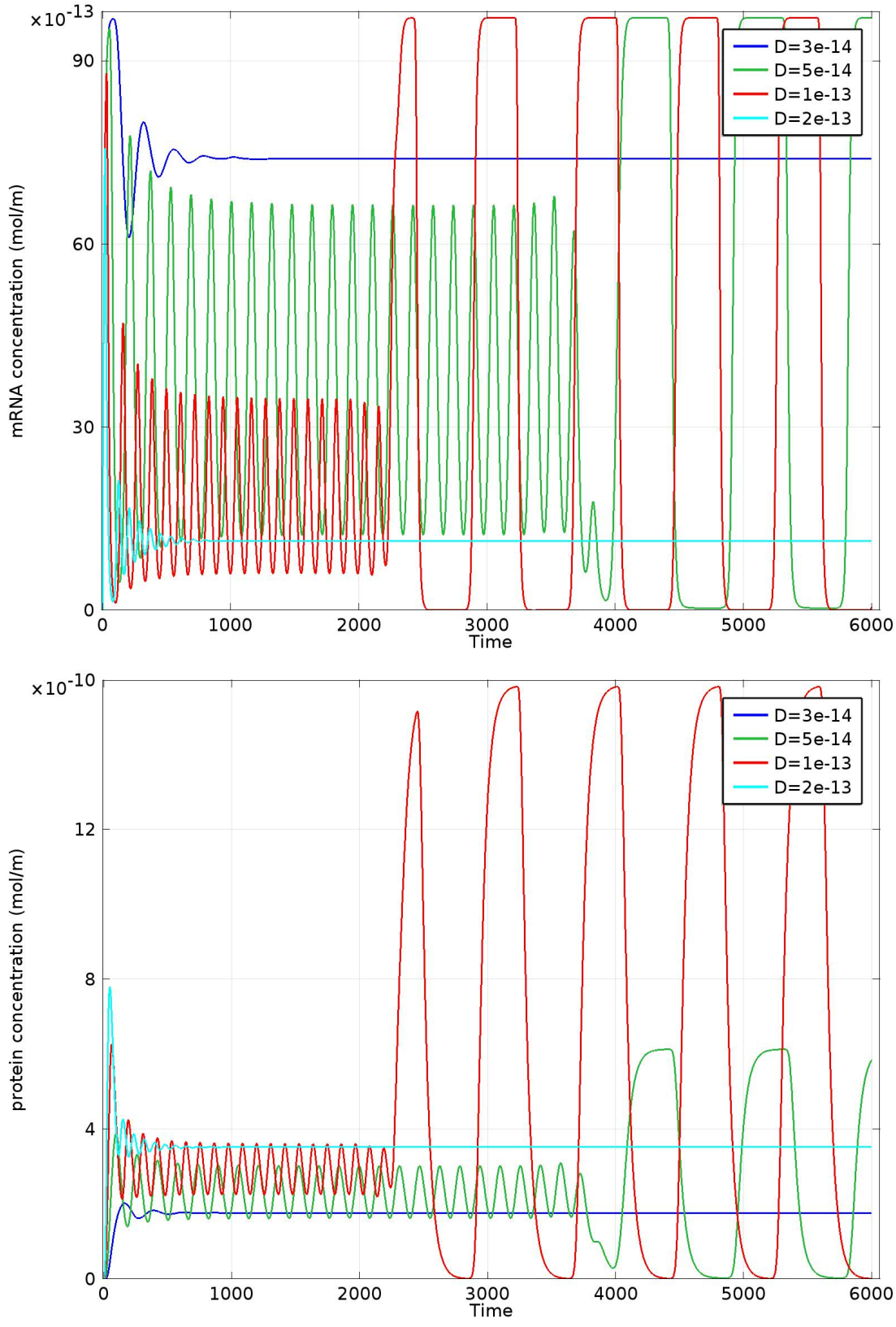

Figure 4.53: *Five-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. A shared gene site at the origin: varying D*

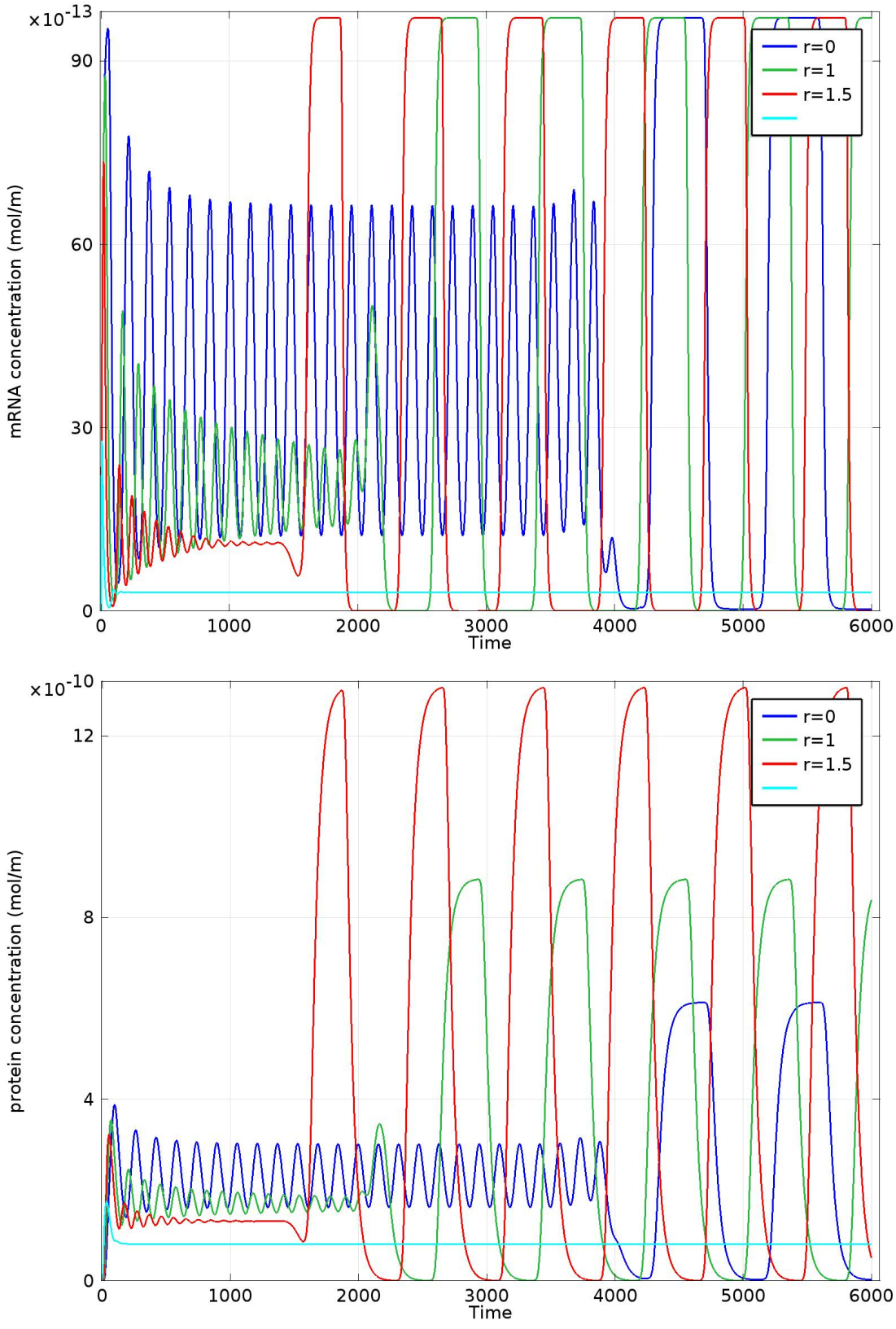

Figure 4.54: *Five-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Varying the location, r, of the shared gene site,* $D = 5 \times 10^{-14}$

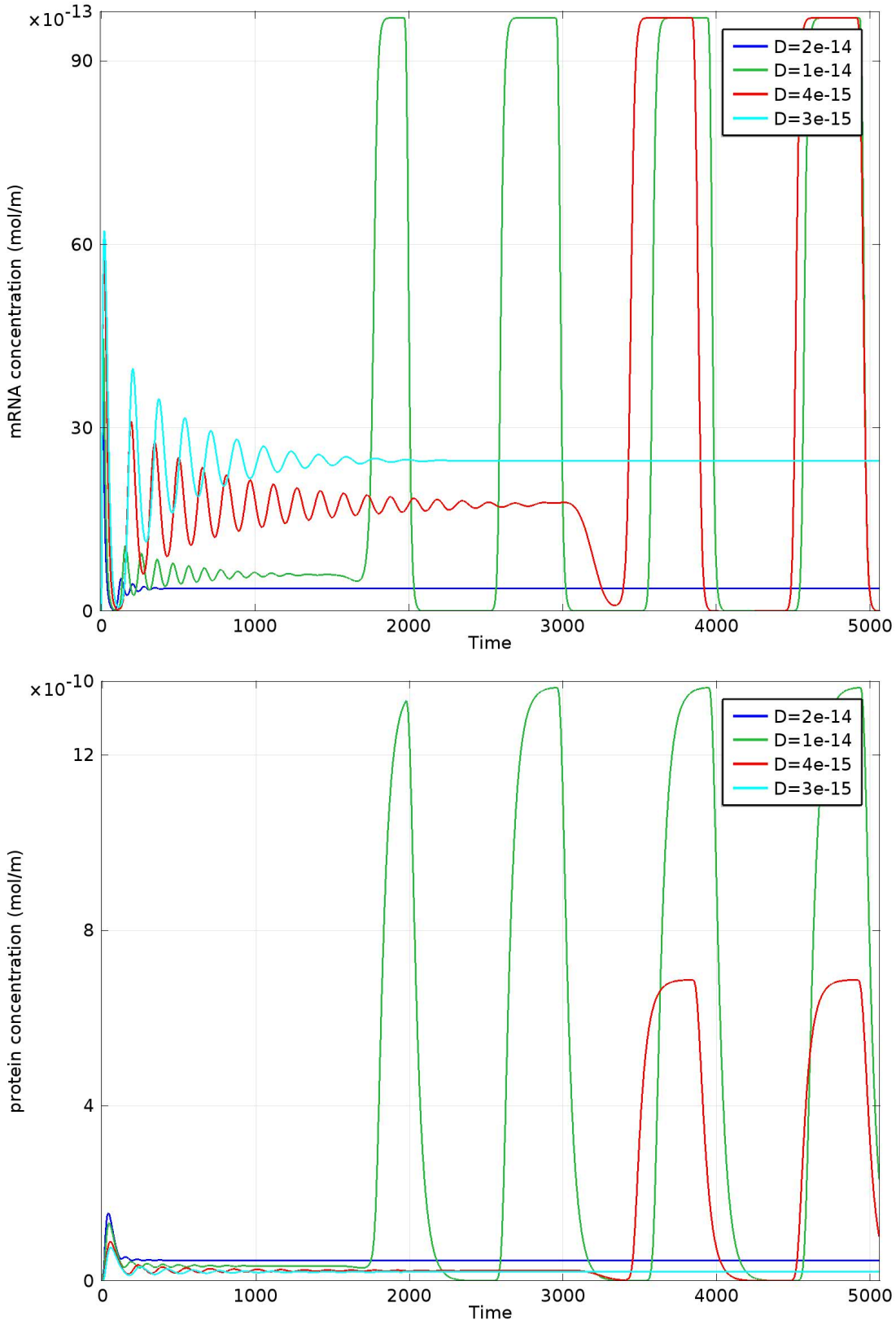

Figure 4.55: *Five-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. A shared gene site near the nuclear membrane, $r = 2.4\mu m$: varying D*

4.6.2 Five individual gene sites

We now consider five individual gene sites existing as a non intersecting cluster at the centre of the nucleus. The location of gene 1, gene 2, gene 3, gene 4 and gene 5 being, $(x_1, y_1) = (0, g_y \mu m)$, $(x_2, y_2) = (g_y \cos(\pi/10), g_y \sin(\pi/10))$, $(x_3, y_3) = (g_y \cos(3\pi/10), -g_y \sin(3\pi/1$ $(x_4, y_4) = (-g_y \cos(3\pi/10), -g_y \sin(3\pi/10))$ and $(x_5, y_5) = (-g_y \cos(\pi/10), g_y \sin(\pi/10))$. $D = 5 \times 10^{-14}$ respectively, where $g_y = 0.86 \mu m$. Thus, they are in a pentagonal configuration of sides of length $1 \mu m$.

We find that the results of the five gene repressilator, with individual gene sites, are similar to that of the three gene repressilator, with individual gene sites. The early and later phase dynamics, transient time of the early phase and change in concentration amplitude are all relatively the same. This still holds after consideration of varying the gene site location and in varying the diffusion coefficient. Hence, under the same parameter values as before and for a similar set of diffusion and gene site location parameter ranges, we find that the transient time of the initial phase of dynamics consist of synchronous spatial-temporal oscillations and the second phase to consist of asynchronous spatial-temporal oscillations.

The change captured between the single shared gene site model to the individual gene sites model also reflects that of the change seen from a single shared gene to individual gene sites in the three gene repressilator model. Thus, the first phase of spatial-temporal oscillations for single shared gene lasts considerably longer than for the five individual origin clustered genes model. Furthermore, the transient time of the first phase of dynamics is decreased further for an asymmetrical gene site cluster at the nuclear membrane. Also, again for changes to the diffusion coefficient, we see that the transient time, period of oscillation, amplitude of concentration, particularly of protein is deeply altered.

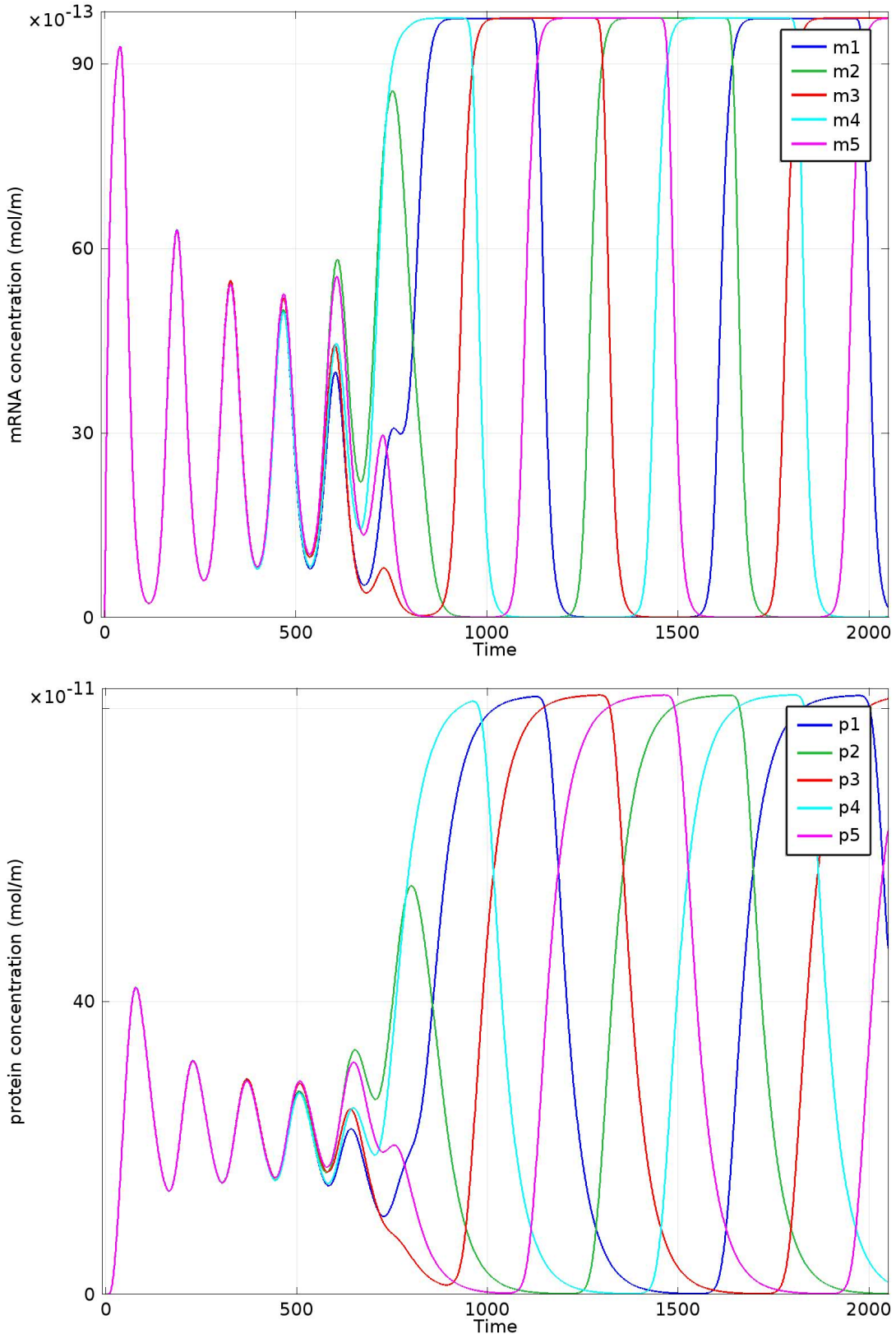

Figure 4.56: *Five-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Individual gene sites clustered together at the origin,* $g_1 = (0, g_y \mu m)$, $g_2 = (g_y \cos(\pi/10), g_y \sin(\pi/10))$, $g_3 = (g_y \cos(3\pi/10), -g_y \sin(3\pi/10))$, $g_4 = (-g_y \cos(3\pi/10), -g_y \sin(3\pi/10))$ *and* $g_5 = (-g_y \cos(\pi/10), g_y \sin(\pi/10))$. $D = 5 \times 10^{-14}$

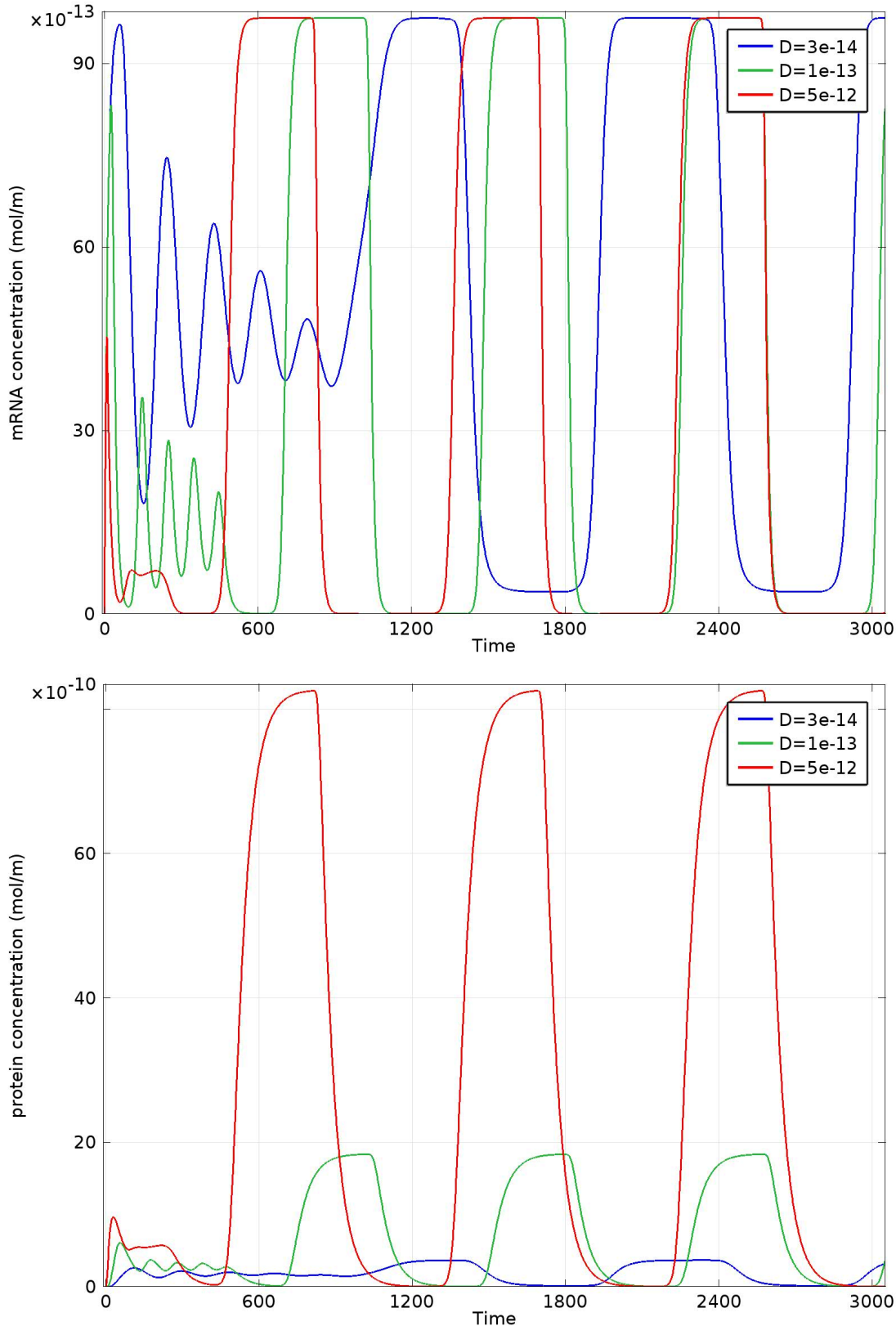

Figure 4.57: *Five-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Individual gene sites clustered together at the origin: varying their diffusion coefficient D,* $g_1 = (0\mu m, g_y\mu m)$, $g_2 = (g_y\cos(\pi/10)\mu m, g_y\sin(\pi/10))\mu m)$, $g_3 = (g_y\cos(3\pi/10)\mu m, -g_y\sin(3\pi/10))\mu m)$, $g_4 = (-g_y\cos(3\pi/10)\mu m, -g_y\sin(3\pi/10))\mu m)$ *and* $g_5 = (-g_y\cos(\pi/10)\mu m, g_y\sin(\pi/10))\mu m)$.

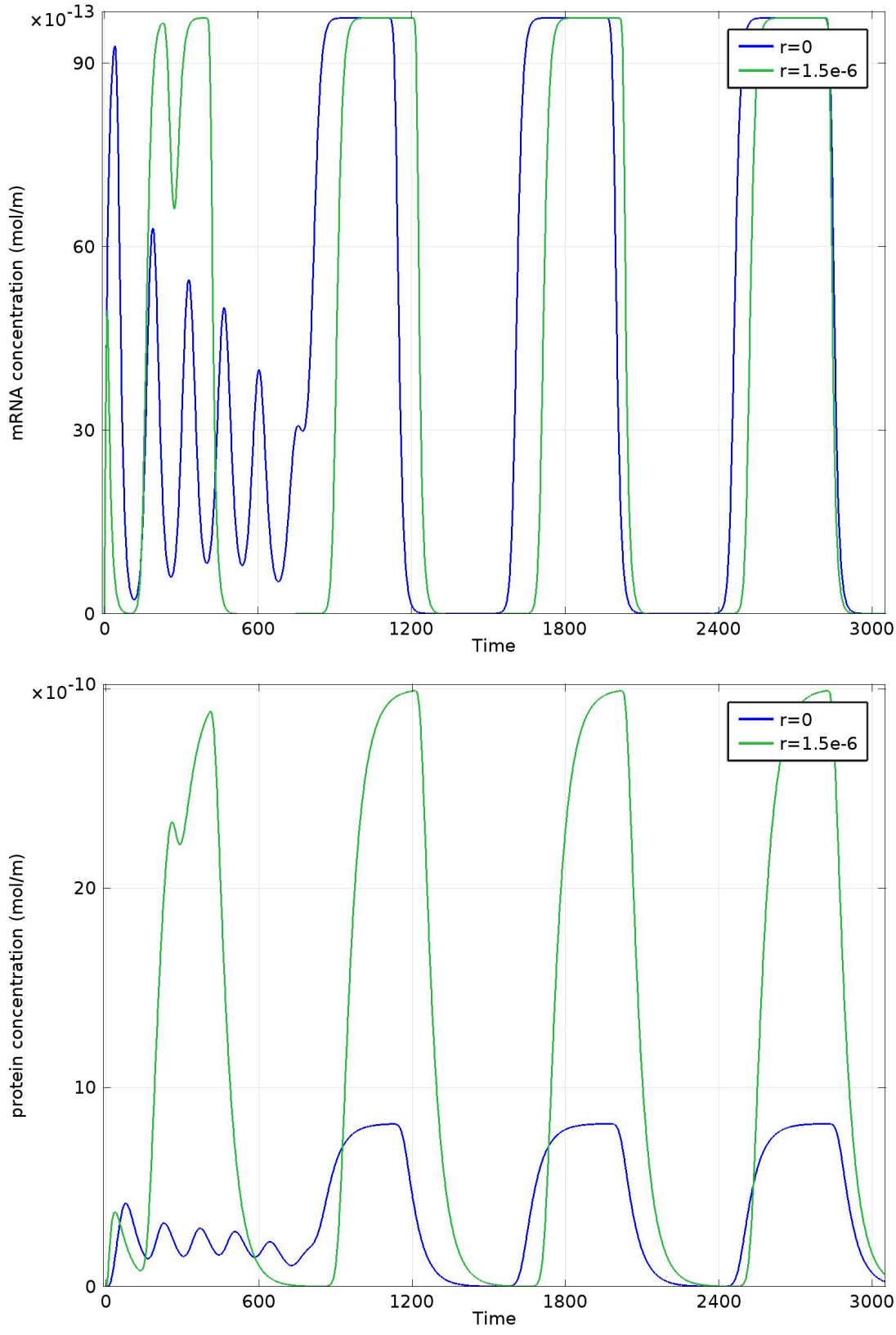

Figure 4.58: *Five-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Individual gene sites clustered together: varying r, their location from the origin,* $g_1 = (0\mu m, g_y + r\mu m)$, $g_2 = (g_y\cos(\pi/10)\mu m, g_y\sin(\pi/10)) + r\mu m)$, $g_3 = (g_y\cos(3\pi/10)\mu m, -g_y\sin(3\pi/10) + r)\mu m)$, $g_4 = (-g_y\cos(3\pi/10)\mu m, -g_y\sin(3\pi/10) + r)\mu m)$ *and* $g_5 = (-g_y\cos(\pi/10)\mu m, g_y\sin(\pi/10) + r)\mu m)$. $D = 5\times10^{-14}$

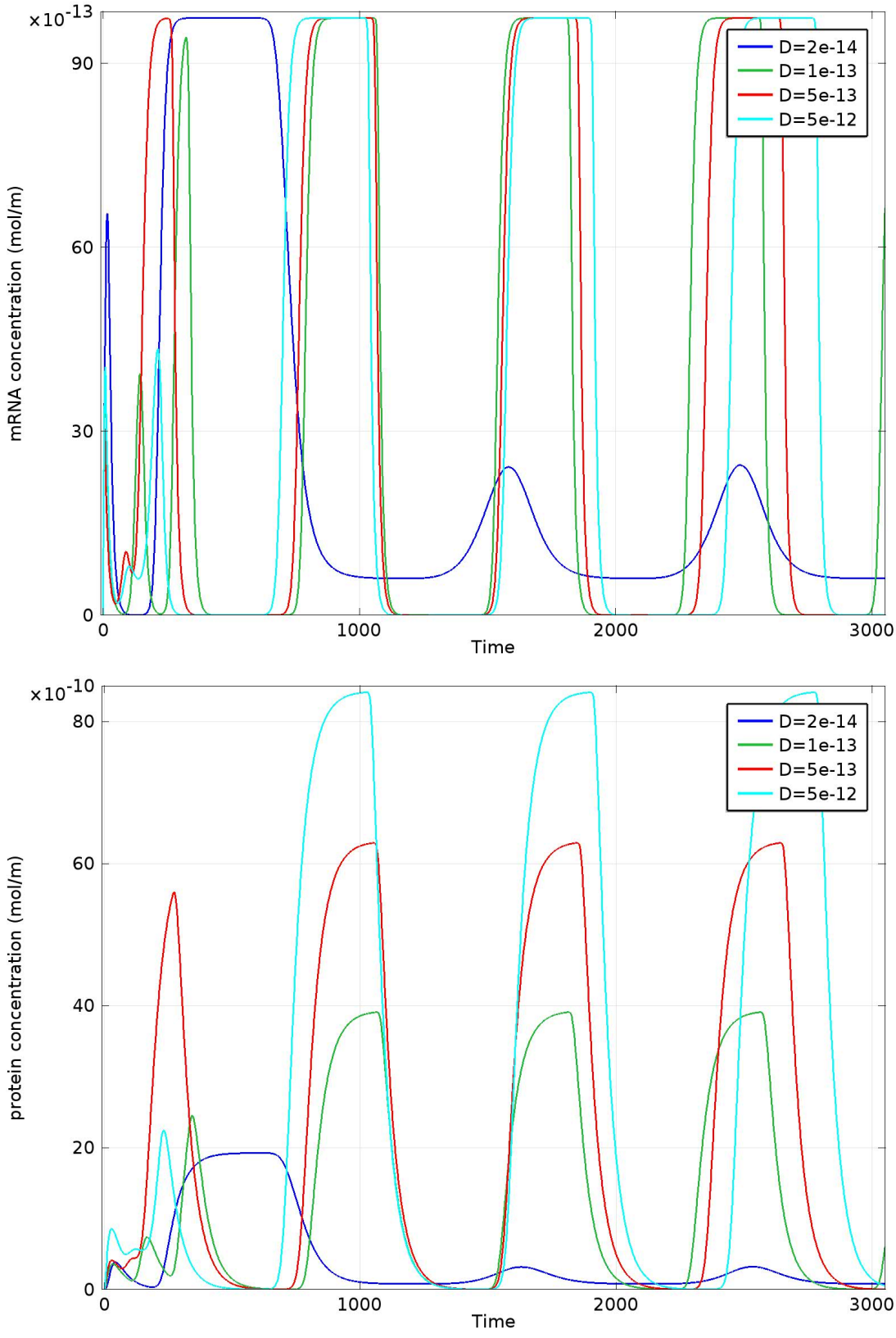

Figure 4.59: *Five-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Individual gene sites clustered together, varying the diffusion coefficient, D.* $g_1 = (0\mu m, g_y + 1.5\mu m)$, $g_2 = (g_y \cos(\pi/10)\mu m, g_y \sin(\pi/10)) + 1.5\mu m)$, $g_3 = (g_y \cos(3\pi/10)\mu m, -g_y \sin(3\pi/10)) + 1.5\mu m)$, $g_4 = (-g_y \cos(3\pi/10)\mu m, -g_y \sin(3\pi/10)) + 1.5\mu m)$ *and* $g_5 = (-g_y \cos(\pi/10)\mu m, g_y \sin(\pi/10)) + 1.5\mu m)$.

4.7 The six gene repressilator

We now consider the six gene repressilator: $n = 6$ of the n-gene repressilator reaction diffusion equations.

4.7.1 A single shared gene site

For the six gene repressilator with a single shared gene site, the spatial temporal dynamics have been found to be oscillatory in the first phase and steady states in the second phase. Again, the period, amplitude of each phase and the transient time of the first phase is heavily dependent on the diffusion coefficient and spatial location of the gene site within the nucleus. The spatial-temporal dynamics of the single shared gene site, six gene repressilator model reflects similar behaviour to that of the single gene, two and four gene rerpressilator models.

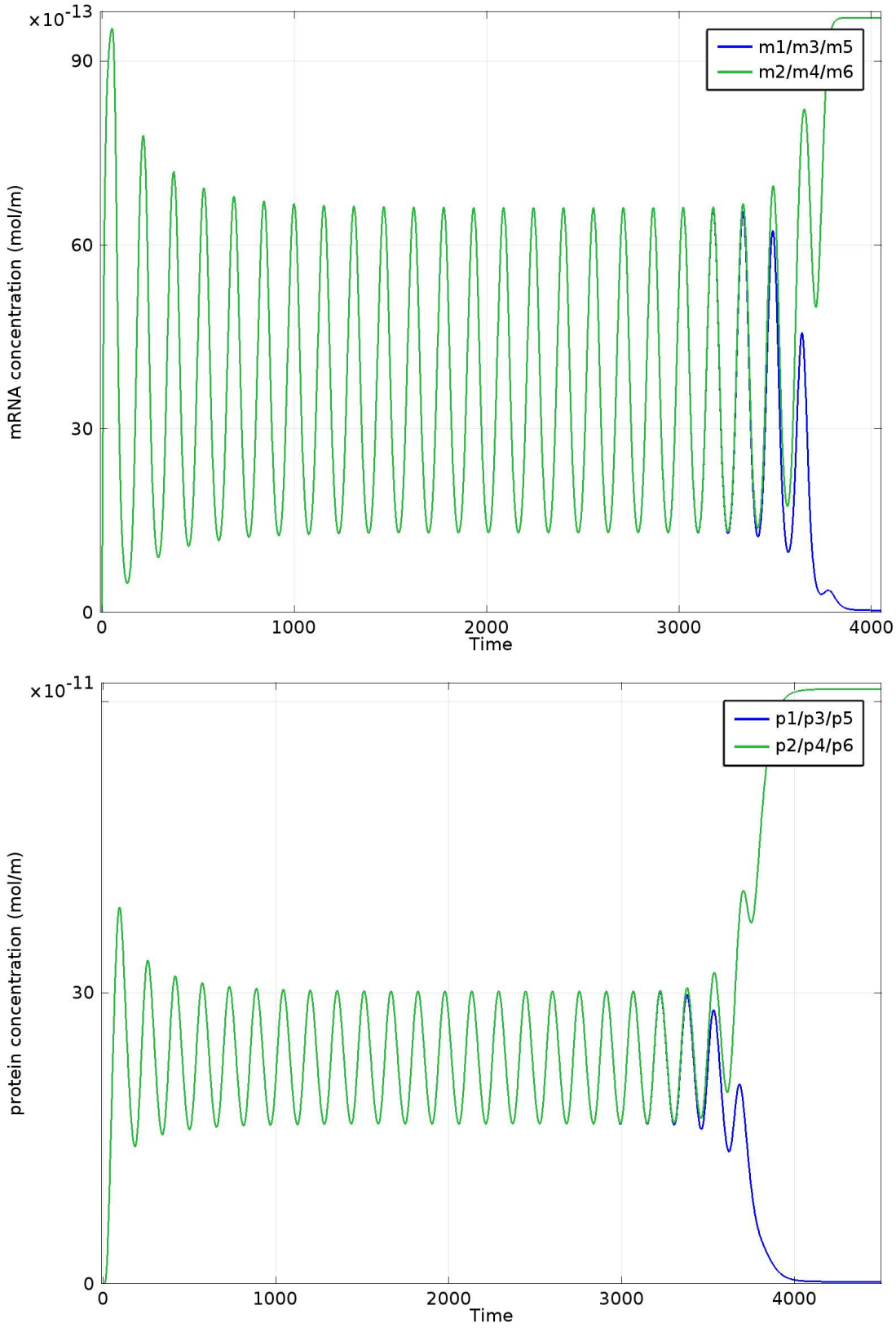

Figure 4.60: *Six-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. A shared gene site at the origin:* $D = 5 \times 10^{-14}$

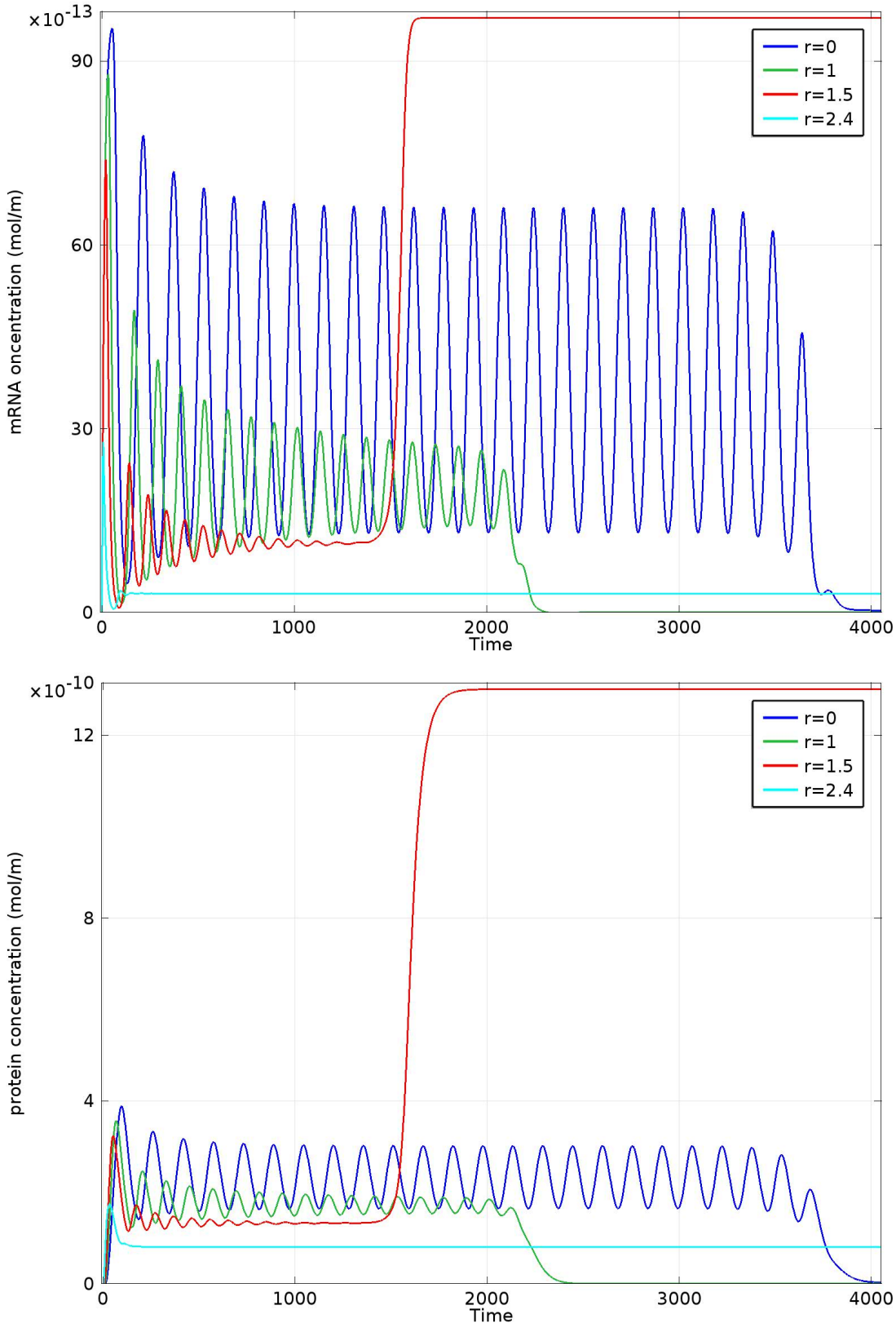

Figure 4.61: *Six-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Varying r i.e. the position of the shared gene site,* $D = 5 \times 10^{-14}$

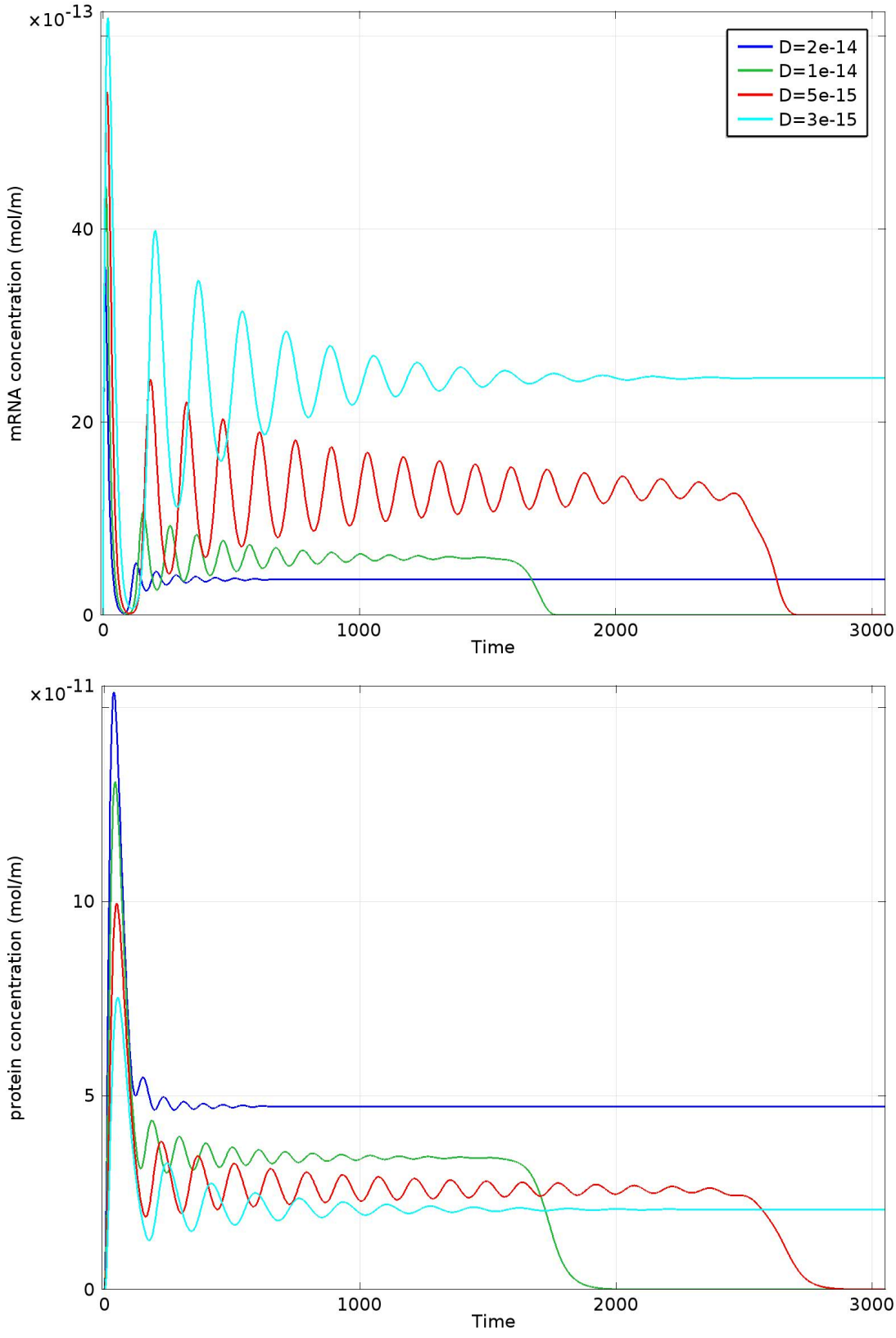

Figure 4.62: *Six-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. A shared gene site, $r = 2.4\mu m$: varying D.*

4.7.2 Six individual gene sites

We now consider six individual gene sites existing as a non intersecting cluster at the centre of the nucleus. The location of gene 1, 2, 3, 4, 5 and 6 being, $(x_1, y_1) = (0\mu m, 1\mu m)$, $(x_2, y_2) = (\cos(\pi/6)\mu m, \sin(\pi/6)\mu m)$, $(x_3, y_3) = (\cos(\pi/6)\mu m, -\sin(\pi/6)\mu m)$, $(x_4, y_4) = (0\mu m, -1\mu m)$, $(x_5, y_5) = (-\cos(\pi/6)\mu m, -\sin(\pi/6)\mu m)$ and $(x_6, y_6) = (-\cos(\pi/6)\mu m, \sin(\pi/6)\mu m)$ respectively. Thus, they are in a hexagonal configuration of sides of length $1\mu m$.

Unlike the two and four (even) gene repressilator, which demonstrated similar dynamics between the shared gene site model and the individual gene site model, in terms of first and second phase of dynamics, the six gene repressilator demonstrates out of phase spatial-temporal oscillations on the second phase of dynamics for individual gene sites, in a cluster at the origin. When we consider varying the diffusion coefficient, we find a small range of D such that the first phase of dynamics are synchronous oscillations and the second phase are asynchronous oscillations: $3^{-14} \leq D \leq 1 \times 10^{-13}$. However, as the diffusion coefficient was increased, the second phase of asynchronous spatial-temporal oscillations changed to steady state solutions.

When we consider moving the gene cluster in the y direction, their is an initial peak in species concentration. An attempt to oscillate is made but then the dynamics retreat to steady state solutions.

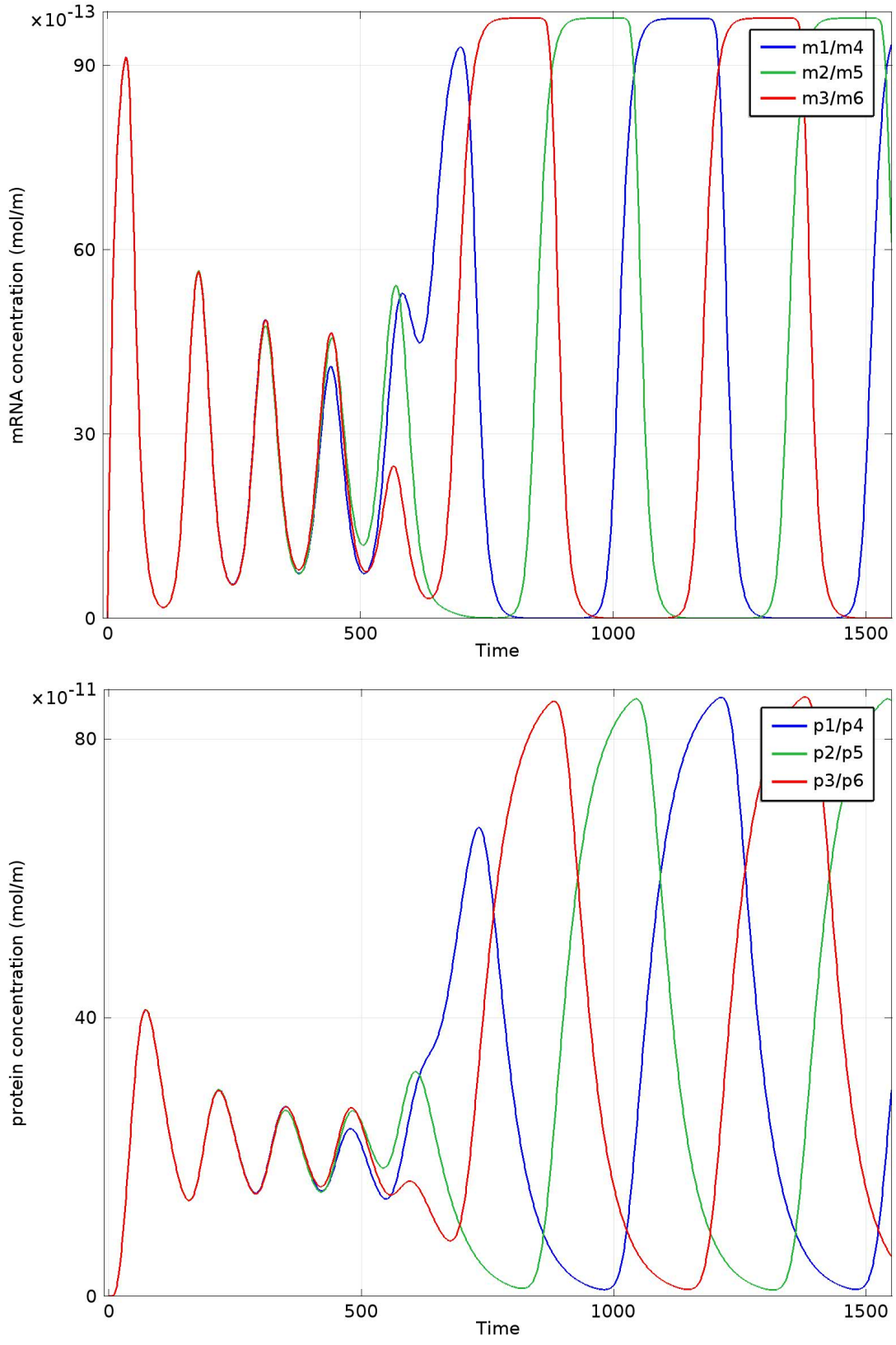

Figure 4.63: *Six-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Individual gene sites, clustered together around the origin:* $g_1 = (0\mu m, 1\mu m), g_2 = (\cos(\pi/6)\mu m, \sin(\pi/6)\mu m), g_3 = (\cos(\pi/6)\mu m, -\sin(\pi/6)\mu m), g_4 = (0\mu m, -1\mu m), g_5 = (-\cos(\pi/6)\mu m, -\sin(\pi/6)\mu m)$ *and* $g_6 = (-\cos(\pi/6)\mu m, \sin(\pi/6)\mu m)$. $D = 5 \times 10^{-14}$

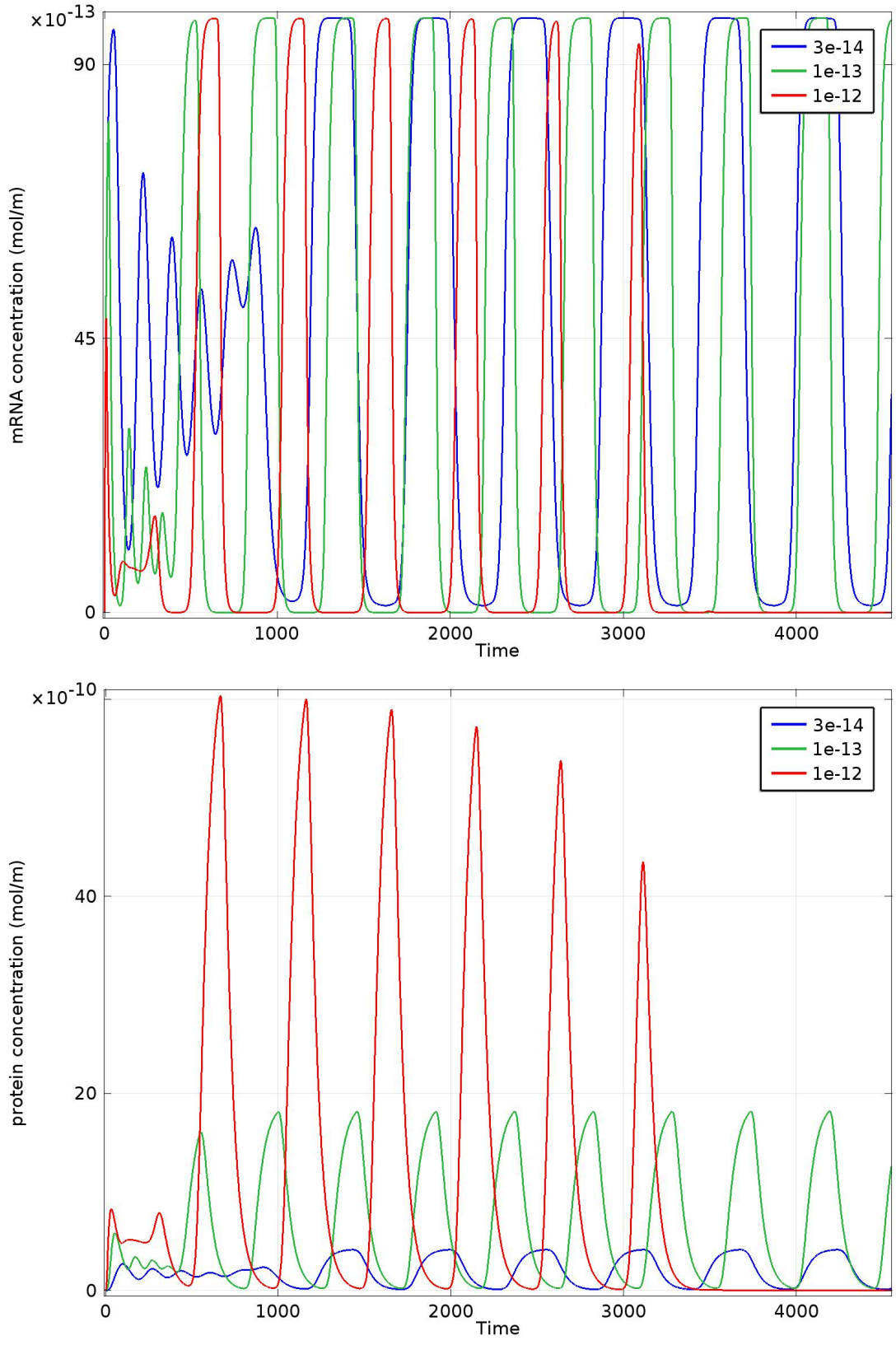

Figure 4.64: *Six-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Individual gene sites, clustered together around the origin:* $g_1 = (0\mu m, 1\mu m), g_2 = (\cos(\pi/6)\mu m, \sin(\pi/6)\mu m), g_3 = (\cos(\pi/6)\mu m, -\sin(\pi/6)\mu m), g_4 = (0\mu m, -1\mu m), g_5 = (-\cos(\pi/6)\mu m, -\sin(\pi/6)\mu m)$ *and* $g_6 = (-\cos(\pi/6)\mu m, \sin(\pi/6)\mu m)$. *Varying the diffusion coefficient D*

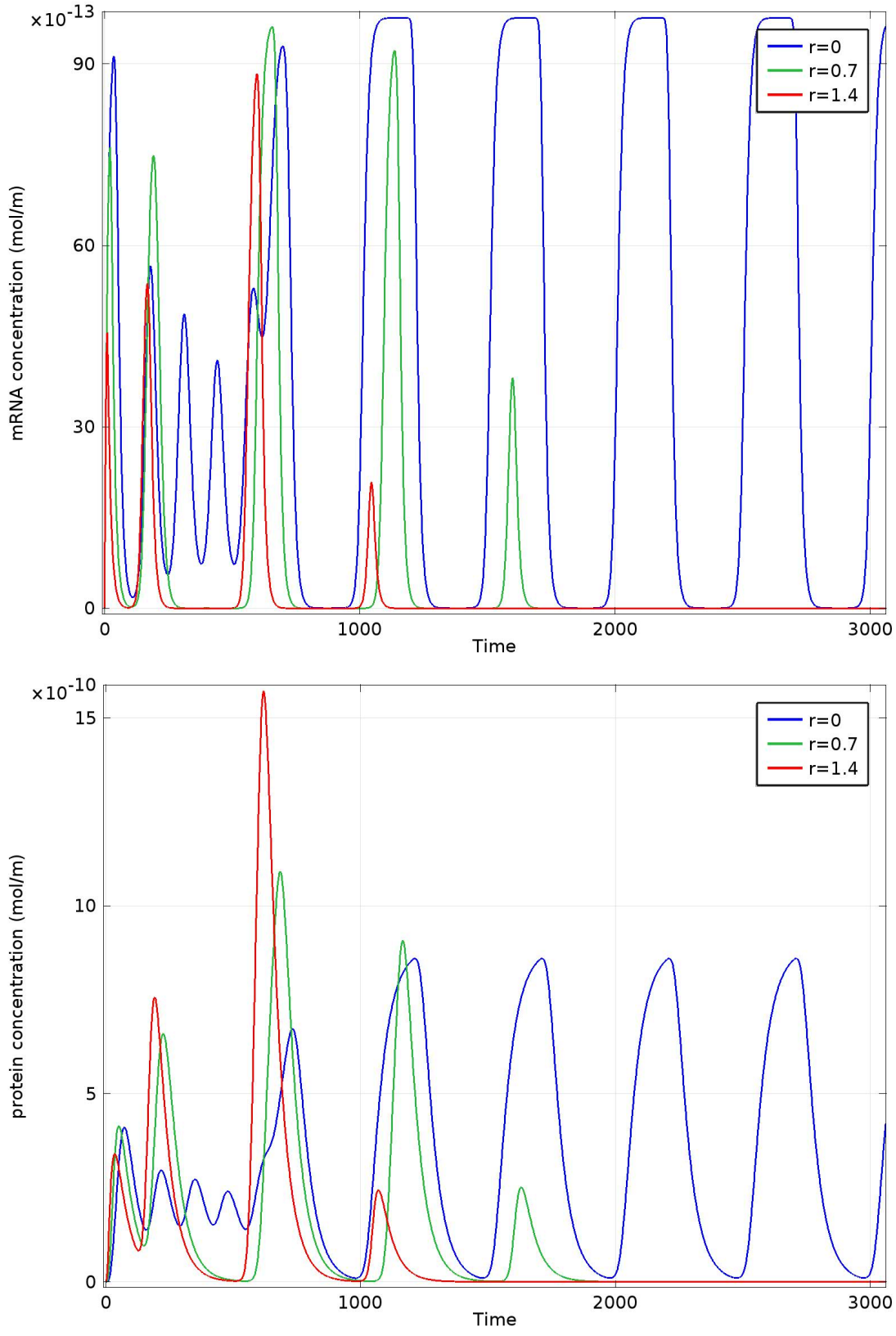

Figure 4.65: *Six-gene repressilator: Plots are of the integration of concentration levels of mRNA (top) and protein (bottom) over the entire cell against time. Individual gene sites, varying r of the gene cluster:* $g_1 = (0\mu m, 1 + r\mu m)$, $g_2 = (\cos(\pi/6)\mu m, \sin(\pi/6) + r\mu m)$, $g_3 = (\cos(\pi/6)\mu m, -\sin(\pi/6) + r\mu m)$, $g_4 = (0\mu m, -1 + r\mu m)$, $g_5 = (-\cos(\pi/6)\mu m, -\sin(\pi/6) + r\mu m)$ *and* $g_6 = (-\cos(\pi/6)\mu m, \sin(\pi/6) + r\mu m)$. $D = 5 \times 10^{-14}$

4.8 Discussion

In this chapter we have considered spatial-temporal models of gene regulatory networks considering both actual (Hes1) and synthetic (repressilator) systems. The construction, modelling and analysis of synthetic GRNs is essential in delineating the underlying mechanisms behind actual GRNs (see, for example, Balagadde et al. 2008, Becskei & Serrano 2000, Chen et al. 2012, Elowitz & Leibler 2000, O'Brien et al. 2012, Purcell et al. 2010). It has been established that GRNs containing negative feedback loops frequently exhibit oscillatory behaviour and mathematical modelling has largely focussed solely on the temporal dynamics using ODE and/or DDE models. Our results show that the dynamics of GRNs can be controlled by spatial components of the PDE model with specific spatial conditions leading to oscillations. This work is in agreement with and generalises previous work by Sturrock et al. (2011*a*,*b*) and Chaplain et al. (2014). It is becoming clear that the spatial components must be considered when modelling GRNs, as they are key to generating periodic behaviour.

More specifically we have investigated the importance of the location of the gene sites within the nucleus. This in turn determines the distance of the gene sites from the cytoplasm, where protein production can occur. We have found that the position of the gene site with respect to the cytoplasm must be optimal in order to achieve oscillations, although the precise optimal ranges will be affected by the size of the diffusion coefficient(s).

It has been demonstrated (Imayoshi & Kageyama 2014, Kageyama et al. 2015, Shimojo et al. 2008) that oscillatory and sustained expression of bHLH transcription factors (such as Hes1) correspond to different states for neural progenitors (self-renewing and fate determining, respectively). Changes to spatial structure provide a realistic mechanism for control of GRNs. Since parameters, diffusion coefficients in particular,

are not well known, we suggest further work is needed to determine parameter regimes more precisely.

We have found that the n-gene repressilator, built solely on negative feedback, preferentially generate oscillations for systems with an odd number of genes. In particular systems with three or five genes are found to oscillate for a wide range of conditions whereas systems with two or four do not. This disparity in behaviour has been found in other models (see Purcell et al. 2010). Our simulations showed that a breaking of symmetry in the cellular domain (positioning of gene sites) lead to an increased tendency for a phase change in dynamics.

However, this generalisation of even numbered genes broke down for $n = 6$. Increasing the number of genes makes the system more robust and likely to oscillate. For an ODE model of multi-gene repressilators, Strelkowa & Barahona (2010) found that increasing the number of genes lead to increased stability of periodic solutions. Our results indicate that this is also true of a comparable PDE system, something which requires further study. If proven, this result may suggest that the complexity due to the number of genes in biologically realistic networks, particularly cascades, are designed in such a way to maximise the likelihood of periodic behaviour while the precise locations of processes may be the key to controlling the timescale of oscillations and any phase change in dynamics needed to begin the next step in cell fate.

It has been shown that molecular crowding can drastically decrease protein diffusion rate (Roosen-Runge et al. 2011). An area of possible research could then look into whether molecular overcrowding in a cell, slowing down molecular diffusion could produce the change in phase of a GRN from oscillatory to steady state. The decrease in diffusion may also arise due to molecular binding, increasing the overall size and hence decreasing its diffusion capacity or due to a decrease in nuclear membrane permeability (Haggie et al. 2004) of which each could arise due to crosstalk with other GRNs (Wang

et al. 2003).

It has been shown that optical deformability is sensitive enough to monitor the subtle changes during the progression of mouse fibroblasts and human breast epithelial cells from normal to cancerous and even metastatic state (Guck et al. 2005). This could possibly change the position of the gene sites and thus possibly change the GRN spatio-temporal dynamics phase i.e. from oscillatory to steady state, changing the fate of the cell.

In biological systems, a phase transition can be defined as the transformation from one phenotype or state to another, where different phenotypes can be mapped to distinct states. For example, the cell cycle is known to have four distinct phases: G1, S, G2 and M phases; cell differentiation contains different stages such as cell proliferation, growth arrest and mature differentiation; and cancer development mainly involves many steps such as mutation, promotion and invasion. Obviously, analysing these biological phase transitions will offer valuable clues for understanding life and its dynamics. Therefore, a fundamental but important question is how to trace the temporal characteristics or dynamics of a biological system during a particular phase transition process (Zeng & Chen 2012). As the cell becomes more cancerous, the shape of the cell may become distorted. Thus, due to this new cellular geometry (or asymmetry, in our model) the second phase of dynamics is reached prematurely. This shows that PDE models of GRNs are better suited to studying the molecular dynamics in such systems than equivalent ODE models.

Our investigation into synthetic networks confirms the importance of spatial modelling of GRNs whilst at the same time has delivered some interesting results which merit further investigation and analysis.

Chapter 5

Spatial-stochastic modelling of the n-gene repressilator

In the previous chapter, the n-gene repressilator, ($n = 1, .., 6$), was modelled as a system of PDEs. For this continuum approach to be valid, we made the assumption that the molecular copy numbers of species was high, such that we could describe the levels as concentrations. However, as discussed earlier, regulator numbers (mRNA and transcription factors) are low and there are only at most two of each gene species. Thus, although the PDE approach can capture spatial effects due to positioning of gene sites and variation to the diffusion coefficient that an ODE approach could not capture, the PDE approach will not capture the effects of stochasticity in a single cell and is not appropriate/valid for systems of low numbers of molecules.

We now consider the more realistic discrete, spatial stochastic approach to our synthetic biology system: the n-gene repressilator, ($n = 1, .., 6$). We begin by first considering the work of Sturrock et al. (2013) where a spatial-stochastic model of the Hes1 GRN: one gene repressilator, was proposed. Each section that follows, we continue with this same approach but incorporate another gene site into the ring repressilator

network each time until we have built up to the six-gene repressilator.

The difference between this spatial-stochastic model and the former PDE model is that we now account for each species being composed of discrete molecules. The ith free promoter is described by an individual species fp_i. The ith occupied promoter, (the product of protein p_{i-1} binding to promoter fp_i), is denoted, op_i.

Gene site species i.e. fp_i and op_i, are confined to their gene site locations, and are thus given a diffusion coefficient of $D_{fp_i} = D_{op_i} = 0$. The processes of transcription, translation and degradation all now have a probability of occurring. Trajectories are sampled from the reaction-diffusion master equation (RDME) as discussed in the general spatial-stochastic approach chapter 3, section 3.3.

The following table describes the biochemical reactions and transportive mechanisms (accounting only for diffusion in this chapter) of the n-gene repressilator, $(n = 1,..,6)$ and states the parameter values used. The ith mRNA species is denoted by m_i, the ith protein species denoted by p_i, the ith free promoter species denoted by fp_i and the ith occupied promoter species is denoted by op_i.

We again choose the diffusion coefficient D to be the same for all diffusive species: $D_{m_i} = D_{p_i} = D, \forall i$.

The PDE model of the n gene repressilator shows that for specific molecular degradation parameters, oscillations will be achieved (at least as an initial transient) provided the relationship between the spatial location of the gene site and diffusion coefficient is satisfied. We will follow a similar protocol here. For set parameter values for all of the reactions in our model we will investigate how dynamics change as we vary the location of the gene site(s) and the diffusion coefficient of the mRNA and protein species.

The instantaneous period of oscillations is estimated using a Morlet continuous time

Cytoplasmic Reaction	Description	Parameter value
$\mathrm{mRNA_i} \xrightarrow{\alpha_{\mathrm{pi}}} \mathrm{mRNA_i} + \mathrm{protein_i}$	Translation of $protein_i$	$\alpha_{pi} = 3min^{-1}$
The ith gene site Reaction	**Description**	**Parameter value**
$fp_i + protein_{i-1} \underset{\mathrm{k_2}}{\overset{\mathrm{k_1}}{\rightleftharpoons}} op_i$	Binding/unbinding of $protein_{i-1}$ to the free ith promoter	$k_1 = 1 \times 10^8 M^{-1} min^{-1}$ $k_2 = 0.05min^{-1}$
$fp_i \xrightarrow{\alpha_{mi}} fp_i + mRNA_i$	basal transcription of $mRNA_i$	$\alpha_{mi} = 3min^{-1}$
$op_i \xrightarrow{\alpha_{mi}/\gamma} op_i + mRNA_i$	Repressed transcription of $mRNA_i$	$\alpha_{mi} = 3min^{-1}, \gamma = 1000$
Global Reactions	**Description**	**Parameter value**
$mRNA_i \xrightarrow{\delta_{mi}} \emptyset$	Degradation of $mRNA_i$	$\delta_{mi} = 0.06min^{-1}$
$protein_i \xrightarrow{\delta_{pi}} \emptyset$	Degradation of $protein_i$	$\delta_{pi} = 0.03min^{-1}$
$\mathrm{S_{ij}} \xrightarrow{\mathrm{d_{jik}}} \mathrm{S_{ik}}$	Molecular diffusion	$d_{jik} = D_i m^2 min^{-1}$

Table 5.1: *The reaction and diffusion processes and their accompanying parameter values used for the n-gene repressilator, discrete spatial stochastic model*

wavelet transform (CWT) as implemented in a MATLAB toolbox called WAVOS, see (Harang & Petzold 2012) for details. Given the highly oscillatory and noisy nature of our trajectories, the use of standard Fourier techniques can lead to inaccurate estimates of the period, as Fourier analysis assumes stationarity of the signal and its basis functions are unbounded in time (Mallat 1998). Wavelets, in contrast, are localised in both time and frequency. This localises the analysis, allowing the changes in signal properties to be tracked over time (Torrence & Compo 1998). Furthermore, we make use of gaussian edge elimination to minimise artefacts in the approximation of the period. We used a sinusoidal function to test the capacity of WAVOS. We found that for a correct period (if any at all) is to be calculated, the time of simulation has to be at least three times as long as the maximum period. An increase on this would lead to more period being displayed on the heat map, due to less Gaussian edge elimination.

The instantaneous period calculated for each time step can vary throughout one single trajectory. For each trajectory the average instantaneous period and corresponding standard deviation was calculated. Calculating the modes was a bit trickier and dependent on parameters chosen.

For each parameter set, at least 100 trajectories were run. Four simulations are shown for each parameter set of D and r chosen.

In order to compare how dynamics change due to a change in diffusion coefficient or gene site location, we will consider averages in instantaneous period, $\bar{T}$ and average and maximum peak value of protein, $\bar{p}_{peak}$.

5.1 The Hes1 GRN: The one gene repressilator

Let D_{s_i} denote the diffusion coefficient of each species, s_i, and let r denote the distance of the gene site from the origin.

Then for each specific value of D and r, four trajectories are shown. This is to try and present a visual idea of the variability in the amplitude and period of the oscillations in molecular numbers.

The discrete, spatial-stochastic model is in agreement with the PDE model as we lower the diffusion coefficient. For a gene site at the origin $r = 0$, we note that there is a minimum $D = D_{syn}$ such that any protein at all can be synthesised. mRNA begins to fail to reach the cytoplasm or does not survive long enough once it reaches the cytoplasm for protein to be synthesised and thus, there is zero protein in the cell for diffusion coefficient $D \leq D_{syn}$. We define D_{syn} as the maximum value of the diffusion coefficient, such that mRNA can not survive long enough to reach and/or reside in the cytoplasm long enough for protein to be synthesised.

$$D \leq D_{syn} = 1 \times 10^{-13} m^2 min^{-1}. \tag{5.1}$$

Additionally as we lower the value of D toward D_{syn} the average and maximum number of mRNA seen in each trajectory increases, due to the decrease in protein production and hence, decrease in negative feedback.

Similarly, we define D_{bind} as the maximum value of D such that protein are unable to survive long enough to diffuse to the gene sites and find and bind to its corresponding target promoter. Thus, for a diffusion coefficient in the range,

$$D_{syn} \leq D \leq D_{bind} = 5 \times 10^{-13} m^2 min^{-1}, \tag{5.2}$$

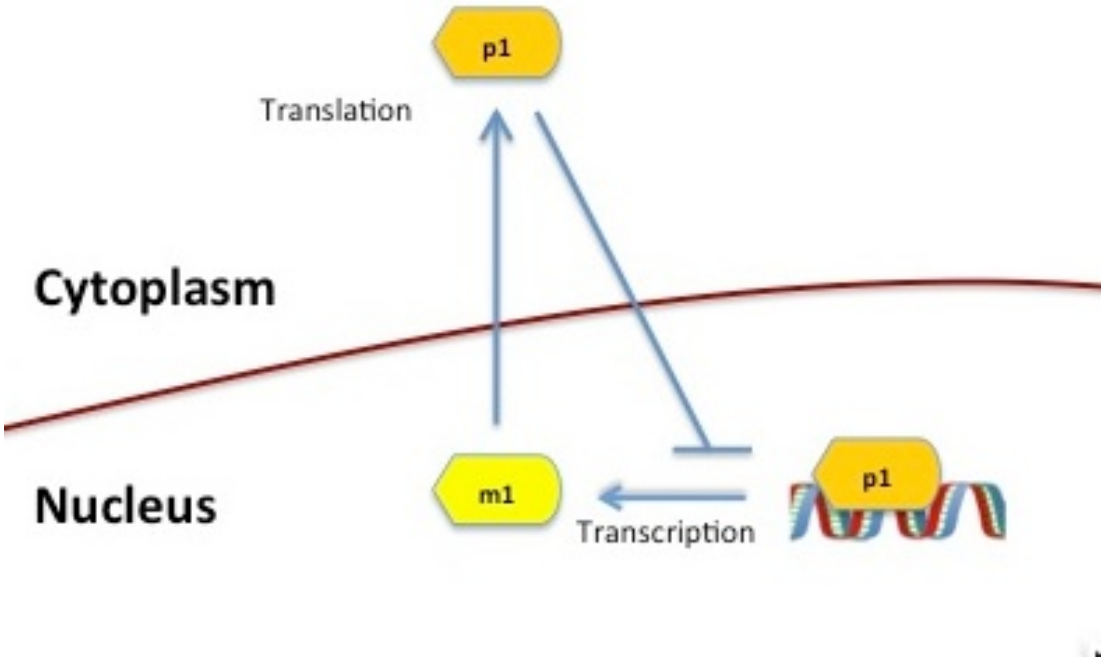

Figure 5.1: *Schematic diagram of the Hes1 GRN: One gene repressilator*

protein and mRNA species are both synthesised but there is no negative feedback.

Therefore, any dynamics in the mRNA and protein levels for that range of D are purely stochastic effects only.

However, if the gene site is located close enough to the nuclear membrane we do see small amplitude spikes of protein produced. Thus, the size of D_{syn} and D_{bind} get smaller as the gene site moves close to the nuclear membrane.

For a diffusion coefficient of $D = 1 \times 10^{-12} m^2 min^{-1} > D_{bind}$, the promoter site of the gene is seen to be occupied at times and thus, the negative feedback in the protein-promoter binding is apparent in the mRNA and protein levels. The first set of four simulations to be shown have $D = 1 \times 10^{-12} m^2 min^{-1}$ and a gene site at the origin, $r = 0$ i.e. $g = (0\mu m, 0\mu m, 0\mu m)$, of the nucleus. The second set of four simulations also have gene site positioned at the origin but are for the faster diffusion regime:

$D = 1 \times 10^{-10} m^2 s^{-1}$.

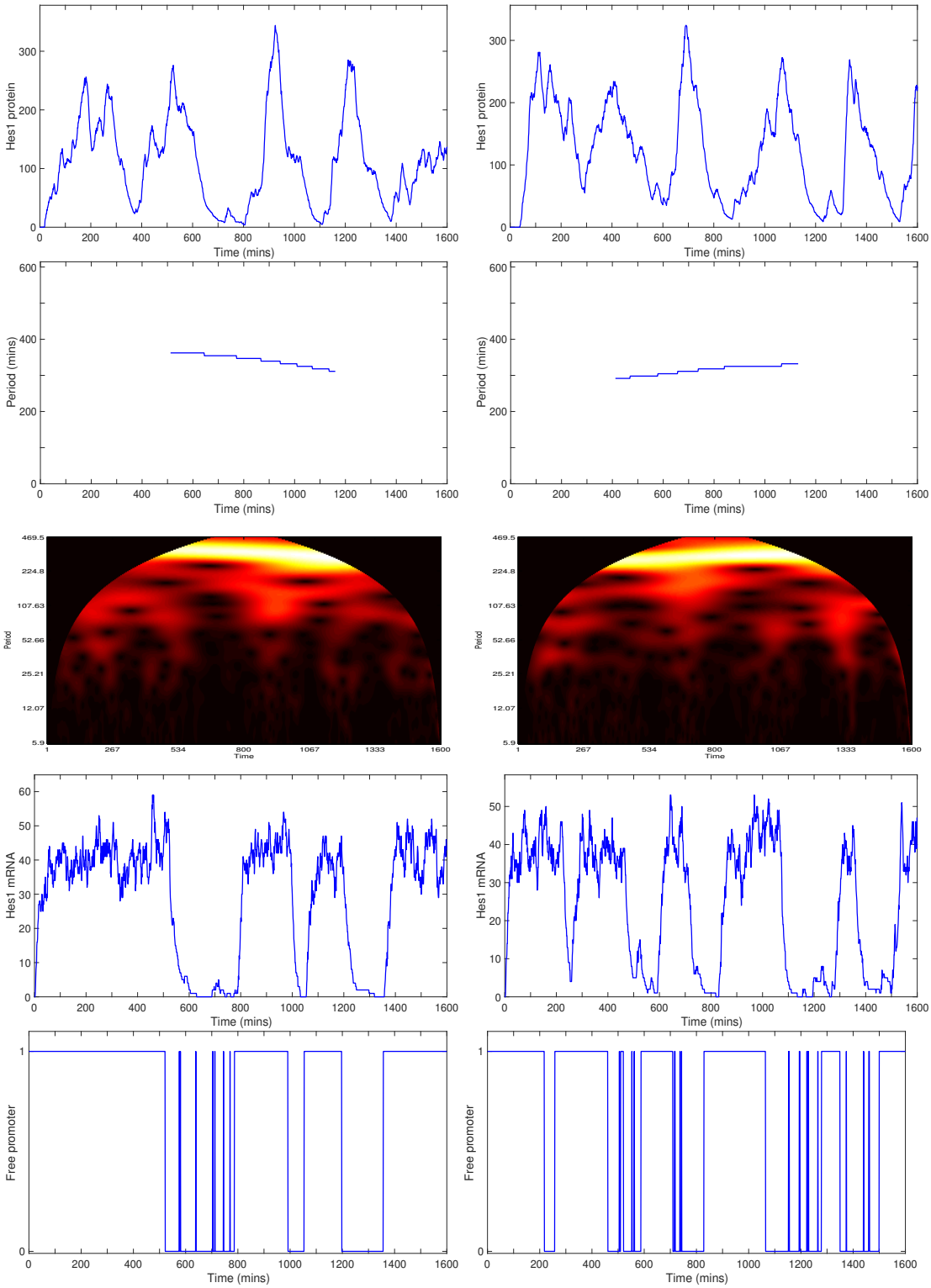

Figure 5.2: *The one gene repressilator with gene site at the origin: $r = 0\mu m$ and diffusion coefficient: $D = 1 \times 10^{-12} m^2 min^{-1}$. Each column represents a realisation. Top to bottom is: total protein number over entire cell, instantaneous period, period heat map, total number of mRNA over entire cell and free promoter time series.*

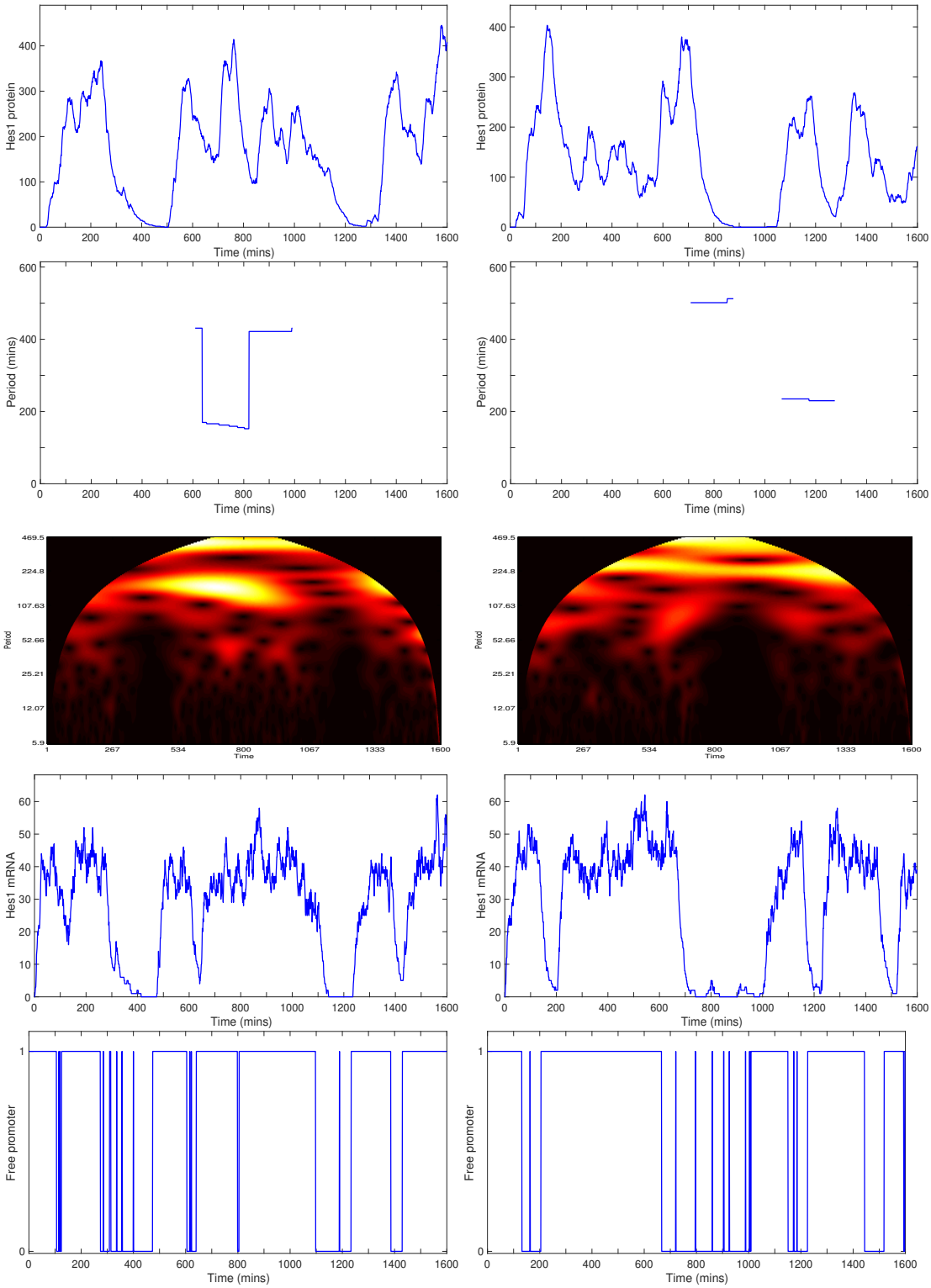

Figure 5.3: *The one gene repressilator with gene site at the origin:* $r = 0\mu m$ *and diffusion coefficient:* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Each column represents a realisation. Top to bottom is: total protein number over entire cell, instantaneous period, period heat map, total number of mRNA over entire cell and free promoter time series.*

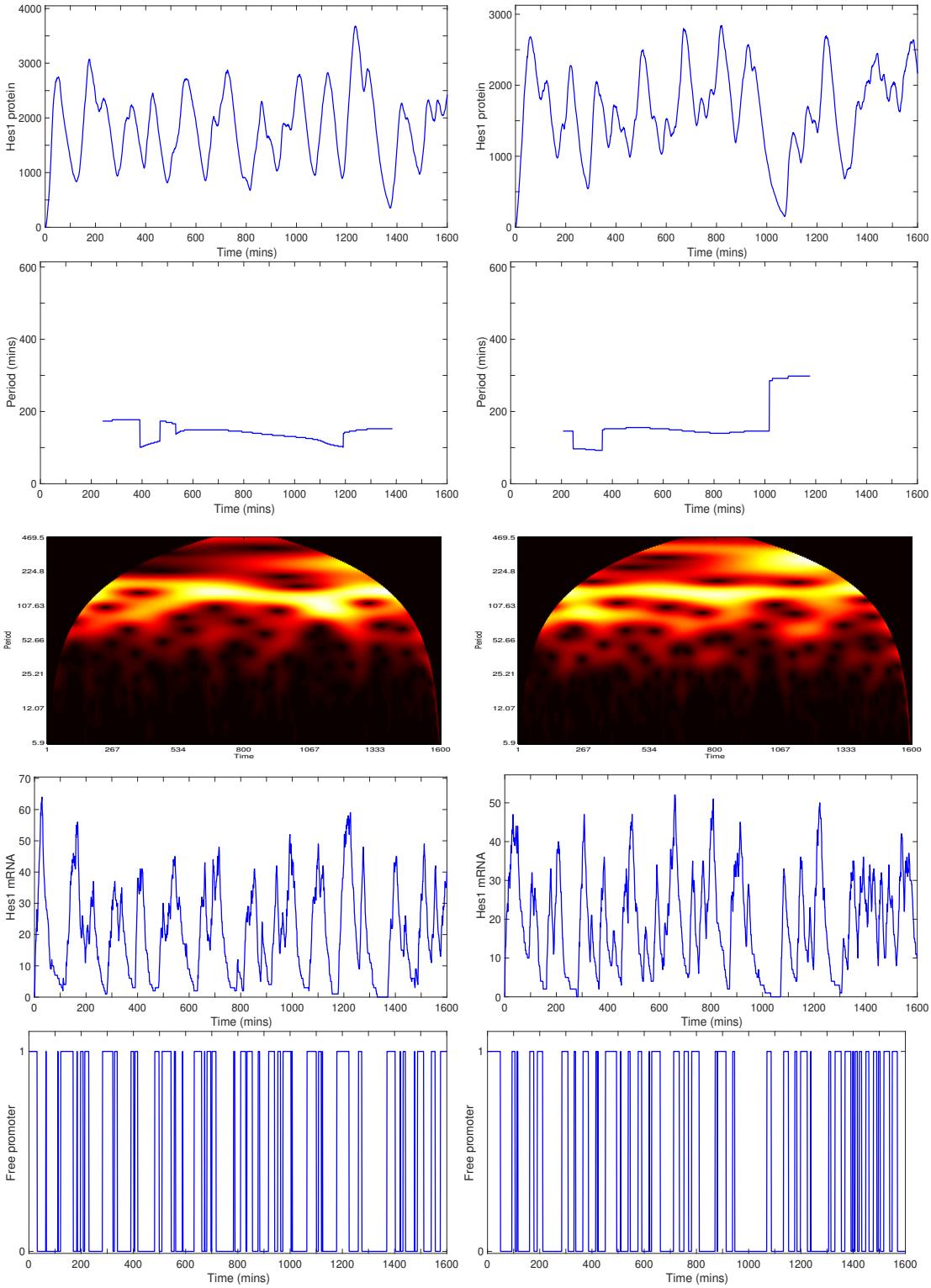

Figure 5.4: *The one gene repressilator with gene site at the origin: $r = 0\mu m$ and diffusion coefficient: $D = 1 \times 10^{-10} m^2 min^{-1}$. Each column represents a realisation. Top to bottom is: total protein number over entire cell, instantaneous period, period heat map, total number of mRNA over entire cell and free promoter time series.*

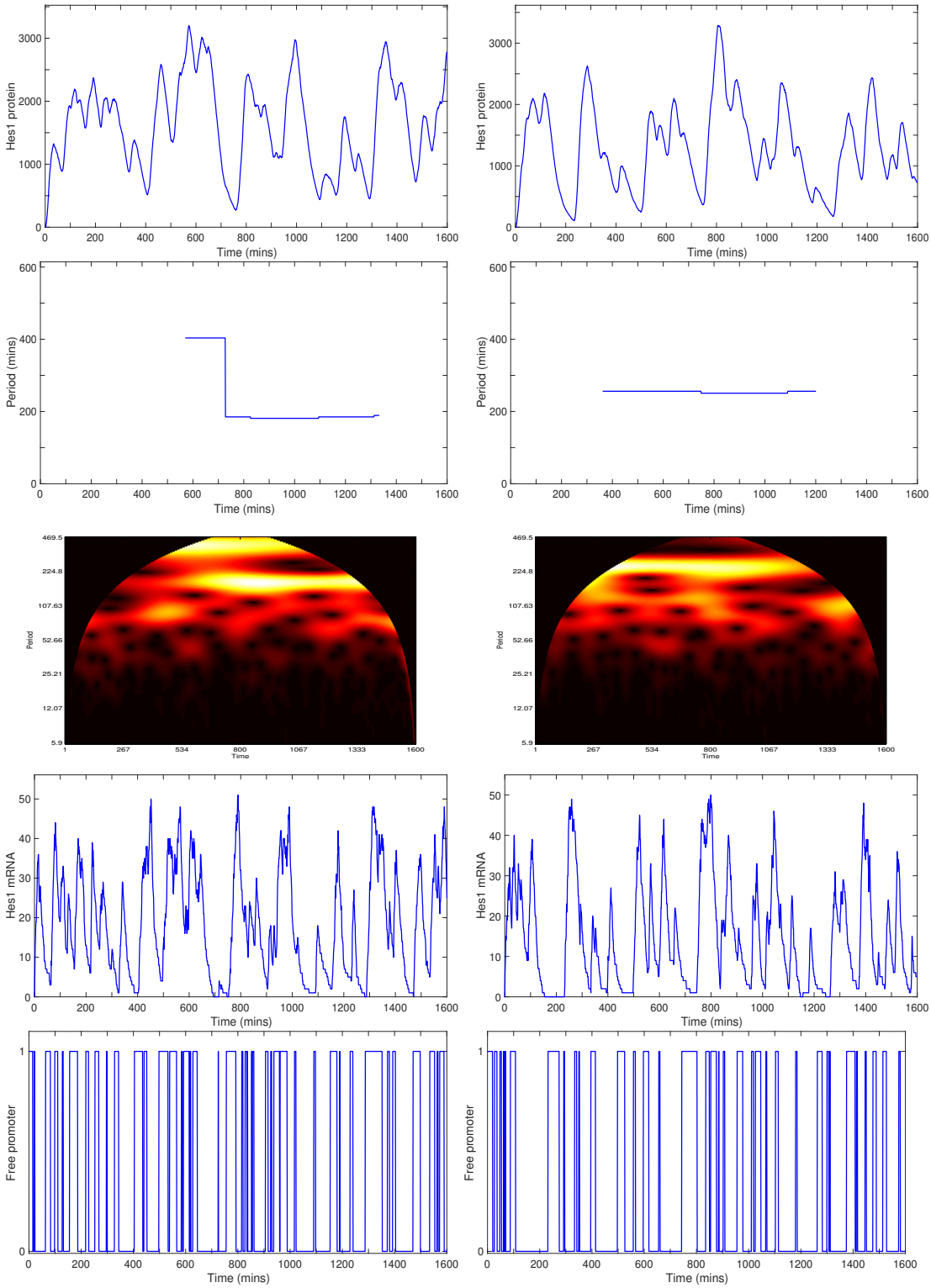

Figure 5.5: *The one gene repressilator with gene site at the origin: $r = 0\mu m$ and diffusion coefficient: $D = 1 \times 10^{-10} m^2 min^{-1}$. Each column represents a realisation. Top to bottom is: total protein number over entire cell, instantaneous period, period heat map, total number of mRNA over entire cell and free promoter time series.*

For $D = 1 \times 10^{-12} m^2 min^{-1}$ the average period for 200 trajectories was $T = 333.1$, ranging between, $93.2 \leq T \leq 558.6$. The average protein number was $\bar{p} = 164.8$, ranging between $83 \leq \bar{p} \leq 234.9$. The average peak value of protein number was $\bar{p}_{peak} = 426.8$, ranging between $253 \leq \bar{p}_{peak} \leq 607$.

For $D = 1 \times 10^{-10} m^2 min^{-1}$ the average period for 200 trajectories was $\bar{T} = 224.5$, ranging between, $224.5 \leq \bar{T} \leq 558.6$. The average protein number was $\bar{p} = 1784$, ranging between $1298 \leq \bar{p} \leq 1784$. The average peak value of protein number was $\bar{p}_{peak} = 3567$, ranging between $2807 \leq \bar{p}_{peak} \leq 4695$.

For a nuclear central gene, the average period of oscillation across trajectories decreased by just short of two hours as the diffusion coefficient, D, was increased by a magnitude of two i.e. from $D = 1 \times 10^{-12} m^2 min^{-1}$ to $D = 1 \times 10^{-10} m^2 min^{-1}$. The average protein number was increased by a magnitude of one due to this increase in the diffusion paramater.

The decrease in average period of oscillation is almost intuitive for an increase in diffusion. However, it can also be explained by the greater amount of protein that can be synthesised and hence reach the gene site and bind to the promoter. Hence, an increase in diffusion leads to a greater negative feedback. However, this increase in potential protein locality around the gene site, leads to a higher level of fluctuation in promoter binding and unbinding. This increase in fluctuations, turning repression on and off more frequently, leads to more complex periodic behaviour, with bi-modal and multi-modal periods detected, see figures 5.6 and 5.7. However, we note, that we need to be cautious of the few modal periods of the smaller diffusion regime that lie on the line of largest period possible that can be detected. To accommodate for this a larger simulation time should be ran. However, we know via experimental data that the period of Hes1 can range between 2 and 5 hours (dependent on the cell type). Thus, at least five of the higher oscillatory periods can fit into the time frame chosen of

1600 minutes. Additionally, the increase in diffusion, enables mRNA to travel to the cytoplasm more readily and hence for protein synthesis, (possibly including protein bursts) to occur at a higher rate. As indeed the average and peak of protein number increases by a magnitude of one due to this increase in the diffusion coefficient.

One could suggest, that this appearance of bi-modal and multi-modal periodic behaviour due to the increase in negative feedback is evidence for more stochastic affects. Although, the range of mean periods is tighter for realisations within the faster diffusion regime, the distribution of periods within each trajectory is more diverse with more bi-modal and multi-modals periods seen. This is rather anti-intuitive but (Huang et al. 2014, Marquez-Lago & Stelling 2010) discuss an increase in negative feedback, increasing stochastic affects. They suggest that this may be due to more frequent protein bursts.

We now consider varying the diffusion for a gene site located near the nuclear membrane: $r = 2.85\mu m$ i.e. $g = (2.85\mu m, 0\mu m, 0\mu m)$. The first set of four simulations are for the slower diffusion regime, $D = 1 \times 10^{-12}$, and the second set of four simulations are for the faster diffusion regime, $D = 1 \times 10^{-10}$.

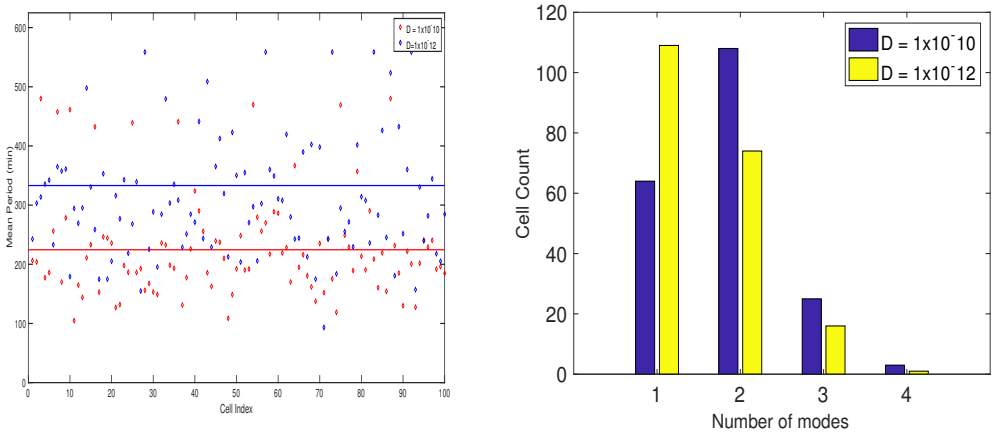

Figure 5.6: *Left: Power spectra of the one-gene repressilator showing average period for 100 trajectories. Red denotes average periods of protein per trajectory under the faster diffusion regime and blue denotes average period per trajectory of the slower diffusion regime. Right: Number of calculated dominant modes appearing in 200 trajectories. Yellow denotes the slower diffusion regime and blue denotes the faster diffusion regime.*

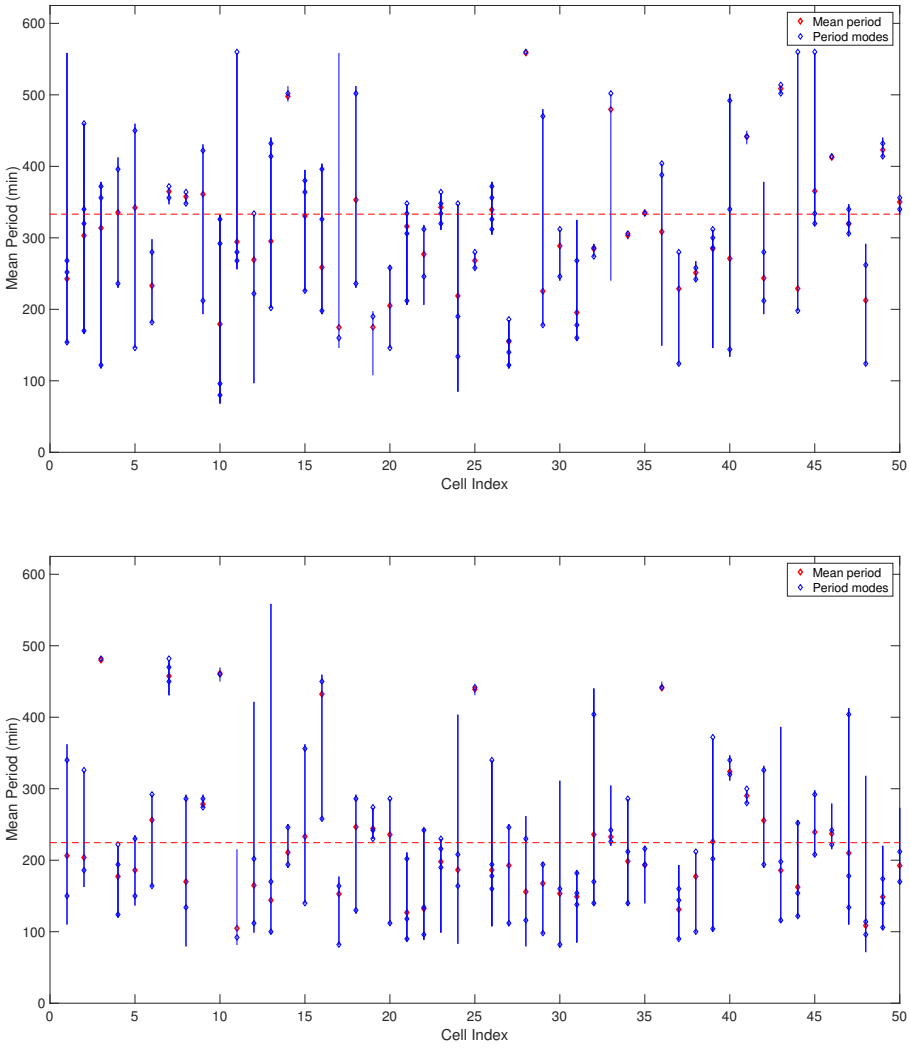

Figure 5.7: *Period data for 50 simulations of the one-gene repressilator model: the red diamonds are the mean periods, the blue lines indicate the range of periods recorded within a trajectory, the blue diamonds indicate the value of the modal period(s). The top image is for the diffusion regime* $D = 1 \times 10^{-12}$ *the bottom plot is for the diffusion regime* $D = 1 \times 10^{-10}$

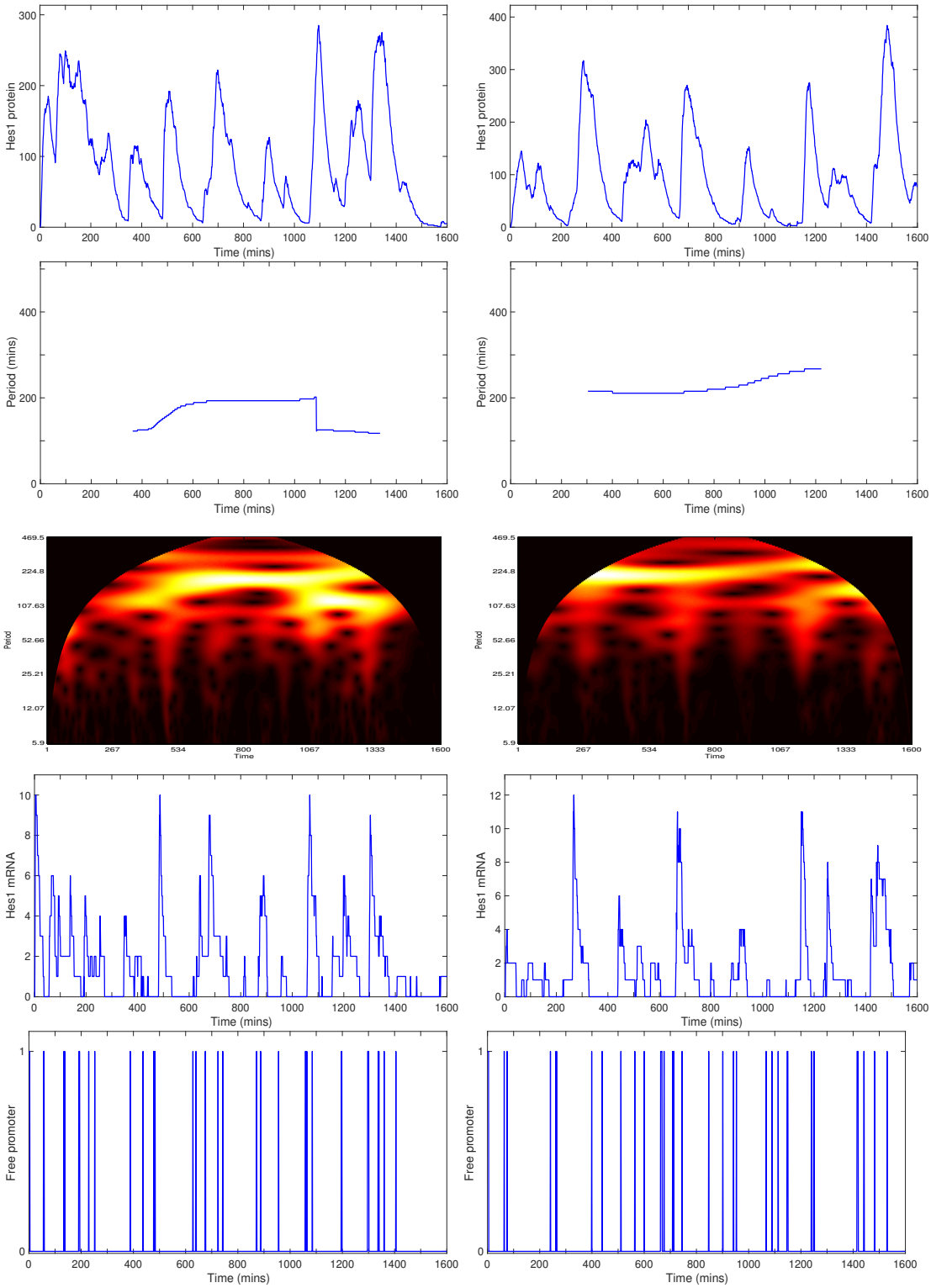

Figure 5.8: *The one gene repressilator with gene site close to the nuclear membrane: $r = 2.85\mu m$, diffusion coefficient: $D = 1 \times 10^{-12} m^2 min^{-1}$: Each column represents a realisation. Top to bottom is: protein, period, period heat map, mRNA and free promoter time series. Translation allowed to occur over entire cytoplasm.*

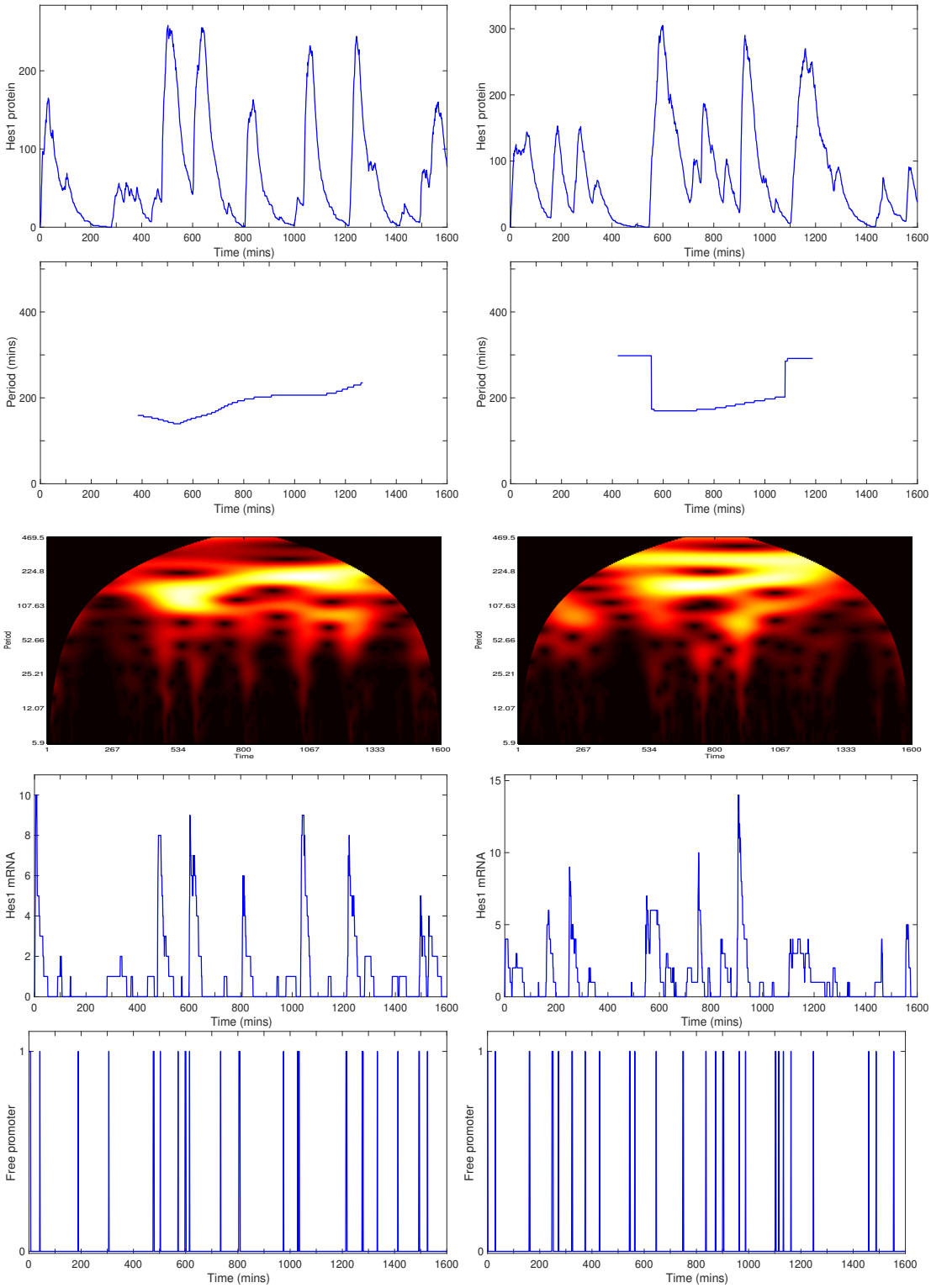

Figure 5.9: *The one gene repressilator with gene site close to the nuclear membrane: $r = 2.85\mu m$, diffusion coefficient: $D = 1 \times 10^{-12} m^2 min^{-1}$: Each column represents a realisation. Top to bottom is: protein, period, period heat map, mRNA and free promoter time series. Translation allowed to occur over entire cytoplasm.*

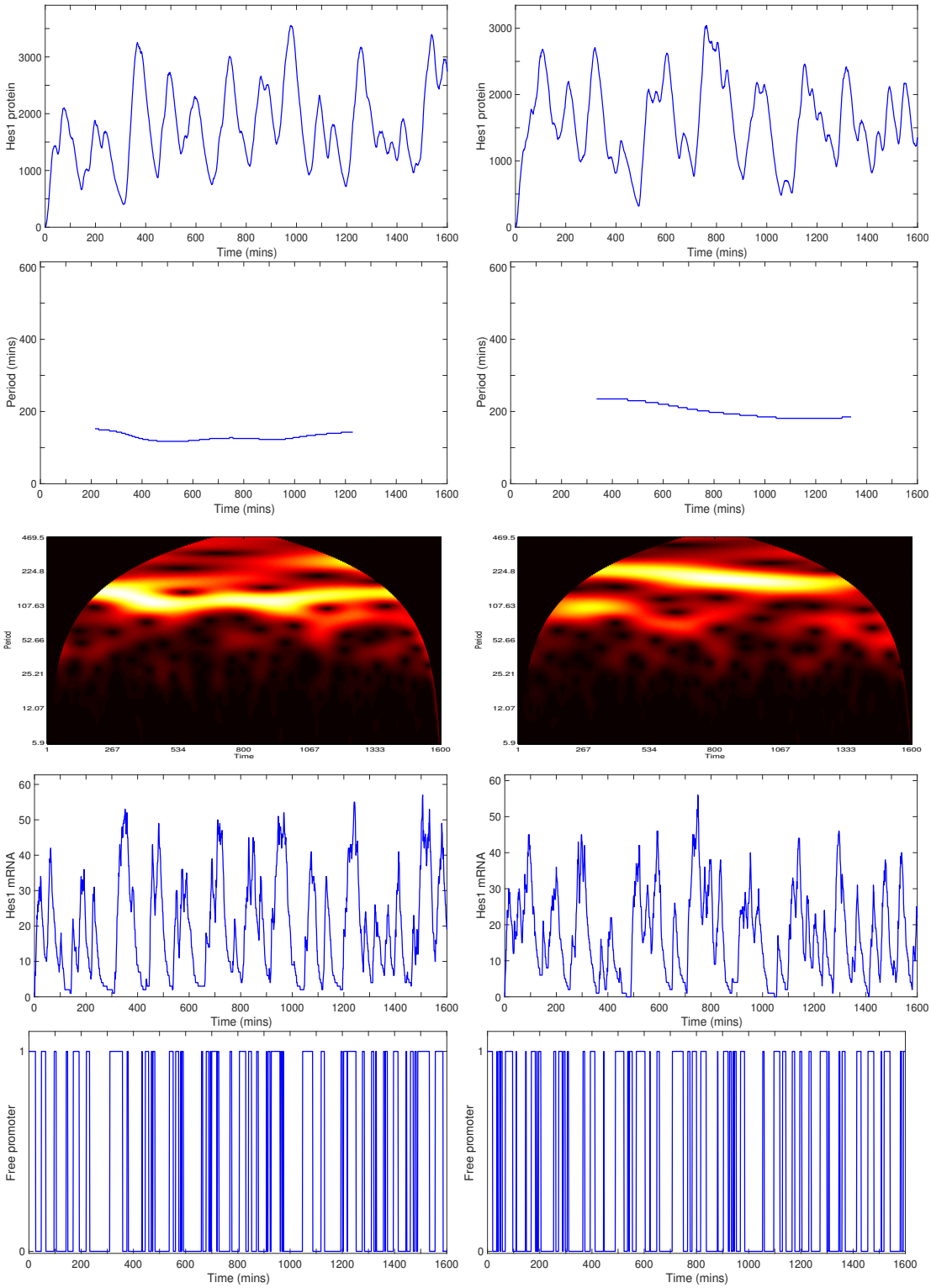

Figure 5.10: *The one gene repressilator with gene site close to the nuclear membrane: $r = 2.85\mu m$, diffusion coefficient: $D = 1 \times 10^{-10} m^2 min^{-1}$: Each column represents a realisation. Top to bottom is: protein, period, period heat map, mRNA and free promoter time series. Translation allowed to occur over entire cytoplasm.*

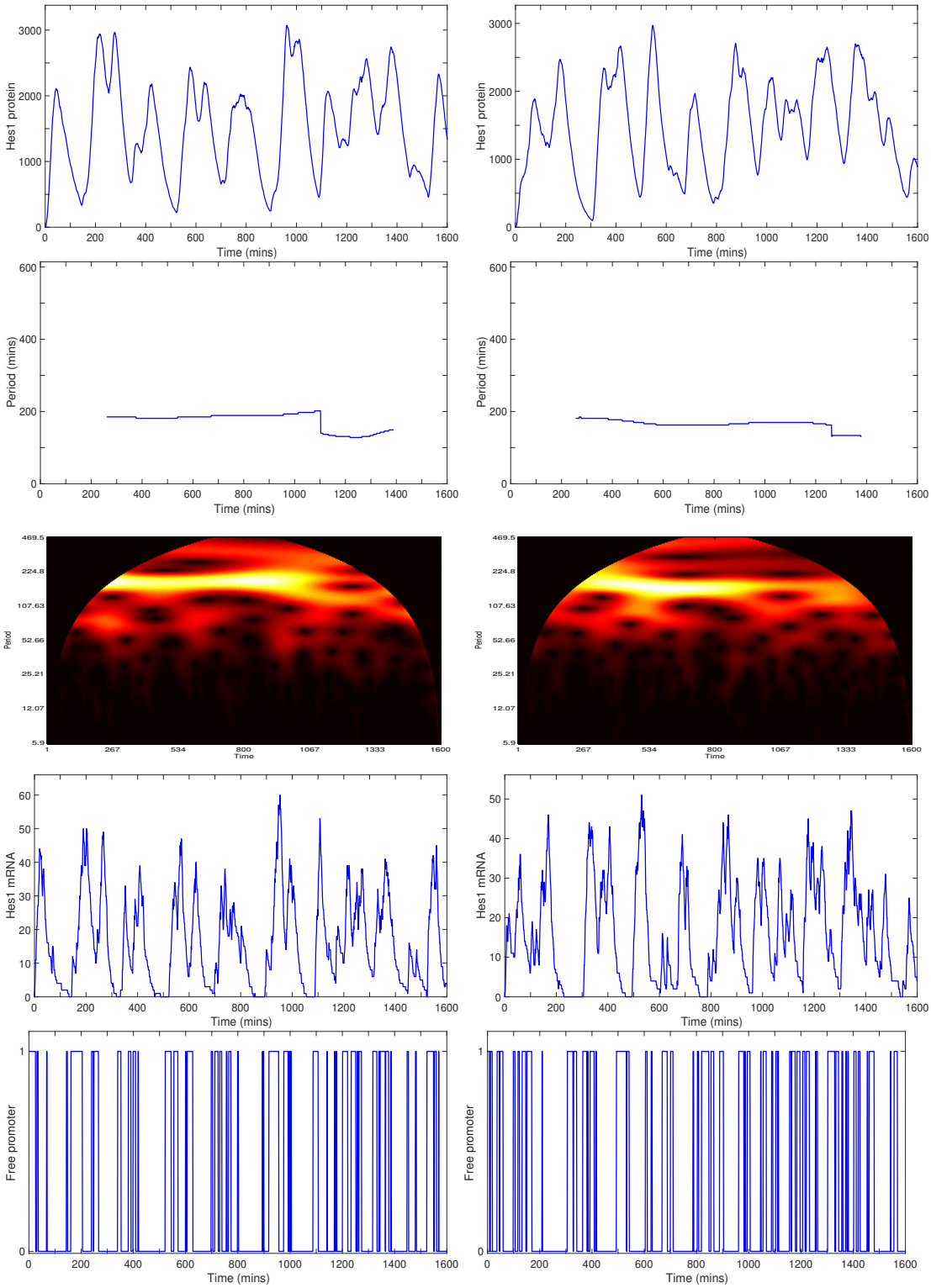

Figure 5.11: *The one gene repressilator with gene site close to the nuclear membrane: $r = 2.85\mu m$, diffusion coefficient: $D = 1 \times 10^{-10} m^2 min^{-1}$: Each column represents a realisation. Top to bottom is: protein, period, period heat map, mRNA and free promoter time series. Translation allowed to occur over entire cytoplasm.*

For $D = 1 \times 10^{-12} m^2 min^{-1}$ the average period for 200 trajectories was $\bar{T} = 258.1$, ranging between, $114.7 \leq \bar{T} \leq 469.7$. The average value of protein number was $\bar{p} = 96.5$, ranging between $56.3 \leq \bar{p} \leq 161.1$. The average peak value of protein number was $\bar{p}_{peak} = 365.3$, ranging between $197 \leq \bar{p}_{peak} \leq 585.6$.

For $D = 1 \times 10^{-10} m^2 min^{-1}$ the average period for 200 trajectories was $\bar{T} = 234$, ranging between, $98 \leq \bar{T} \leq 558.6$. The average value of protein number was $\bar{p} = 1589$, ranging between $1142 \leq p_{av} \leq 1986$. The average peak value of protein number was $\bar{p}_{peak} = 3323$, ranging between $2319 \leq \bar{p}_{peak} \leq 4256$.

For a gene site located near the nuclear membrane, the difference between the mean periods recorded across 200 trajectories as we increase the diffusion parameter by a magnitude of two, is not as significant as we observed for the origin located gene. Although, we note that the faster diffusion regime still exhibits, on average, a somewhat, slightly faster frequency of oscillation. However, there now appears to be a slightly larger spread in the mean periods across trajectories for the faster diffusion regime for a gene site closer to the nuclear membrane than in the origin located gene site case. We now observe bi-modal and multi-modal periods more consistently for both diffusion regimes, see figure 5.12 and 5.13. The increase in the amount of bi-modal and multi-modal periodic behaviour seen for trajectories in the smaller diffusion regime can again be put down to a higher level fluctuation of protein-promoter binding, thus, generating a stronger negative feedback.

Protein levels increase by a magnitude of greater than one this time, compared to just one in the origin located gene site case. The protein levels for the slower diffusion regime slightly decrease as the promoter is moved closer to the nuclear membrane. This is because when the promoter location is very close to the nuclear membrane the promoter is occupied more often than it is free. The consequential effect is far lower levels of mRNA, due to the increased repression by the more frequently protein-bound

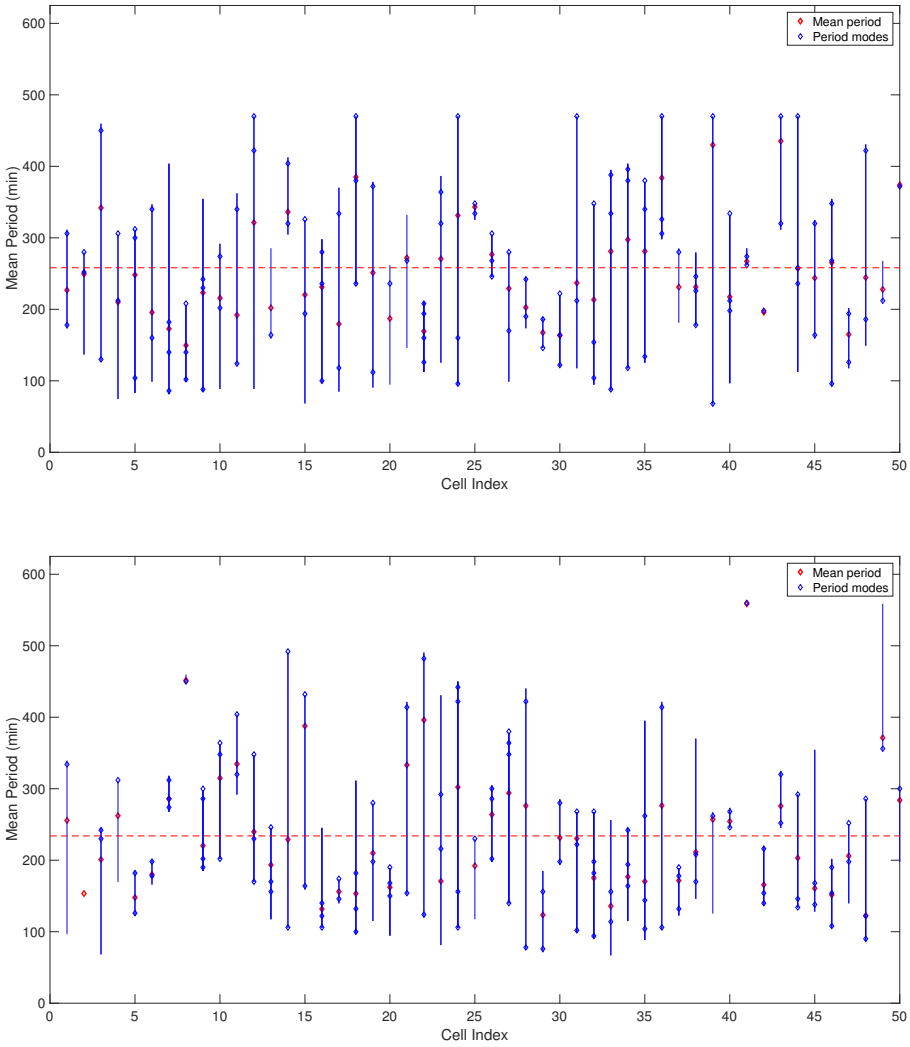

Figure 5.12: *Period data for 50 simulations of the one-gene repressilator model: the red diamonds are the mean periods, the blue lines indicate the range of periods recorded within a trajectory, the blue diamonds indicate the value of the modal period(s). The top image is for the diffusion regime* $D = 1 \times 10^{-12}$ *the bottom plot is for the diffusion regime* $D = 1 \times 10^{-10}$

promoter.

In the PDE model, there is a lower and upper limit of the diffusion coefficient, such that oscillations can be achieved. As the diffusion coefficient is increased past the oscillatory range of D, both the mRNA and protein settle to a steady state. mRNA reaches a smaller level in its steady state concentration as D is increased and protein reacher as higher level in its steady state concentration as D is increased. In the spatial-stochastic regime, there is a lower limit on D but there is no upper limit. A proposal we will make is that the PDE approximation also breaks down here because the stochastic regime is no longer unimodal. The case for the lower diffusion regime, with gene site close to the nuclear membrane, where a higher level of bimodal periodic behaviour is seen, this is equivalent to a faster diffusion regime with gene site at the origin.

The results for the spatial-stochastic model are in agreement with the delay stochastic model mentioned in section 3.2.2.1 Bratsum et al. (2005). Where delay could represent the time required to reach gene site/cytoplasm. However, we have a broader view as it is not only diffusion we take into account, for the delay, but also the spatial position of the gene site within the nucleus of the heterogenous cell.

Figure 5.14 gives a summary of how average protein number and period vary with the diffusion parameter and gene site position.

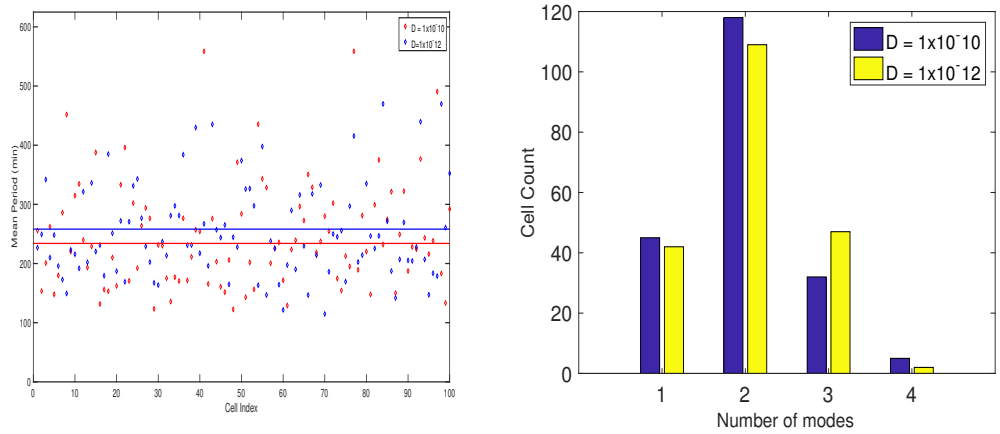

Figure 5.13: *Left: Period data of the one-gene repressilator showing the average period per trajectory for 100 trajectories. Red diamonds denotes average periods of protein level oscillations per trajectory under the faster diffusion regime and blue diamonds denote average period per trajectory of the slower diffusion regime. Right: Number of calculated dominant modes appearing in 200 trajectories. Yellow denotes the slower diffusion regime and blue denotes the faster diffusion regime.*

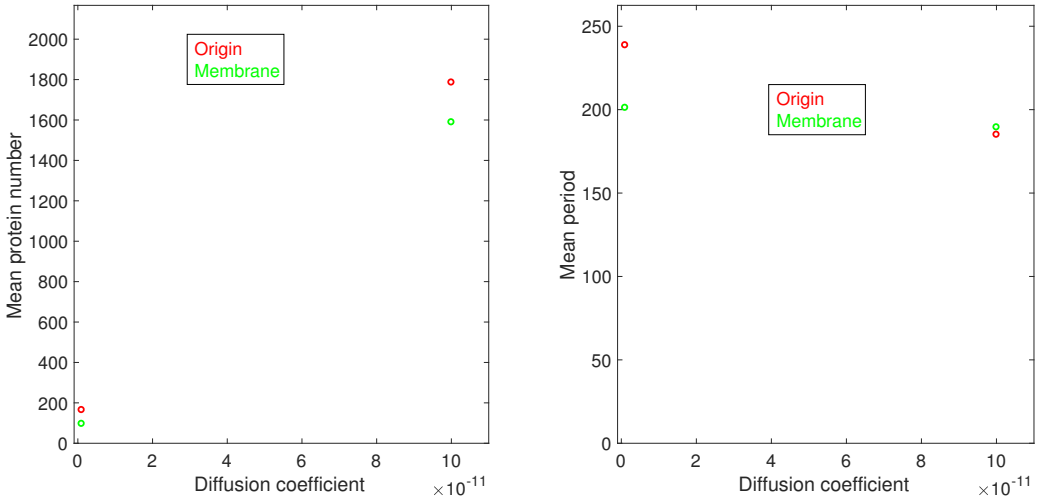

Figure 5.14: *Left: Plot showing the average protein number calculated over 250 trajectories, against the diffusion parameter for low and high diffusion regime. Red is when gene site located at origin, blue is when the gene sites is located close to the nuclear membrane.*

5.2 The two gene repressilator

We now consider the two gene repressilator, ($n = 2$). See figure 5.15 for a schematic of the reaction-diffusion processes considered. We first consider the gene site locations to be $g_1 = (x_1, y_1, z_1) = (0.5\mu m, 0\mu m, 0\mu m)$ and $g_2 = (x_2, y_2) = (-0.5\mu m, 0\mu m, 0\mu m)$ for gene 1 and gene 2 respectively. We term this gene site configuration a cluster at the centre of the nucleus. The first set of three simulations are for the smaller diffusion regime, $D = 1 \times 10^{-12}$ and the second set of three simulations are for the faster diffusion regime, $D = 1 \times 10^{-10}$.

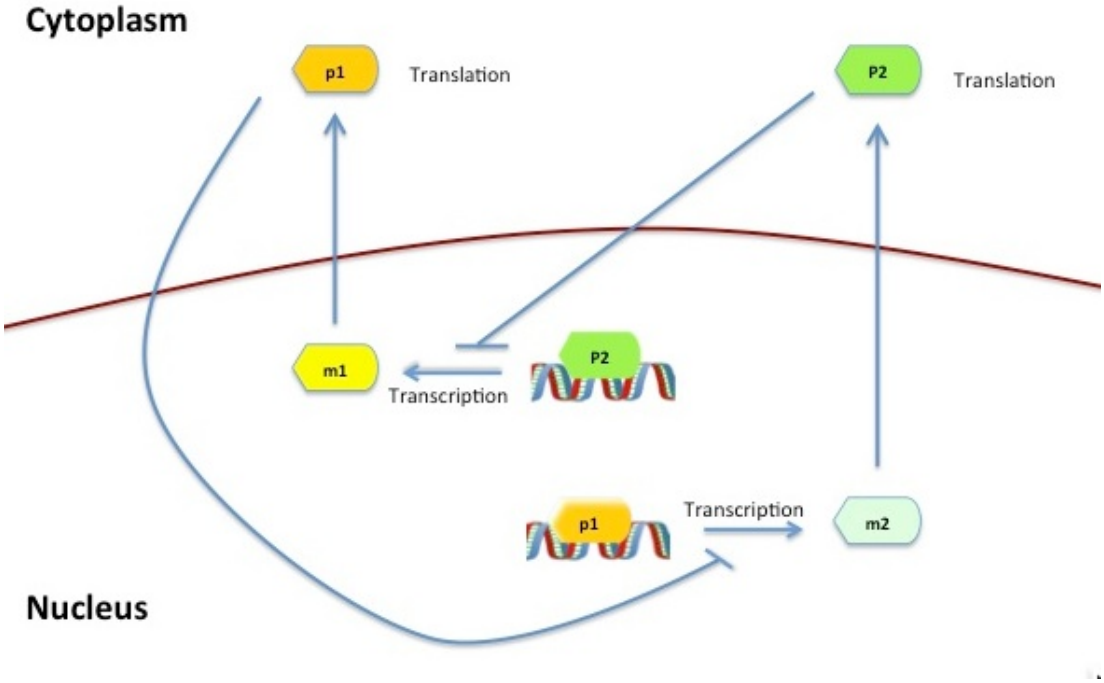

Figure 5.15: *Schematic diagram of the two gene repressilator, $n = 2$. Protein 1 can bind to gene 2 and repress transcription of mRNA 2. Similarly protein 2 can bind to gene 1 and repress transcription of mRNA 1. When mRNA is transcribed is transcribed it is free to diffuse from the nucleus into the cytoplasm. In the cytoplasm protein 1 is translated due to mRNA 1 and protein 2 is translated due to mRNA 2. Once protein are synthesised it is free to diffuse from the cytoplasm to the nucleus. Once a protein finds its target gene, it has the capacity to bind and hence, repress transcription of mRNA.*

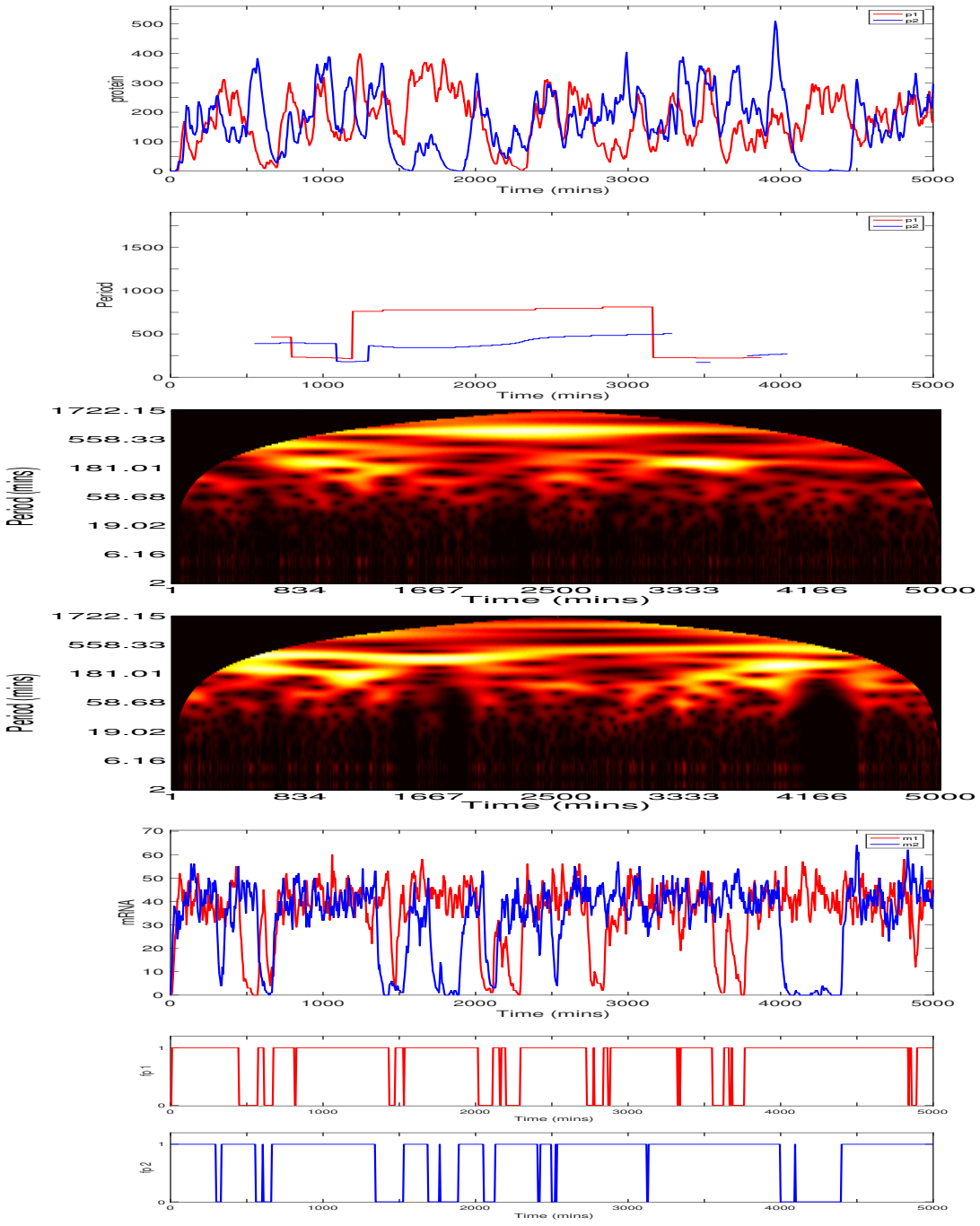

Figure 5.16: *The two gene repressilator with individual gene sites clustered at the centre of the nucleus:* $g_1 = (0.5\mu m, 0\mu m, 0\mu m), g_2 = (-0.5\mu m, 0\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

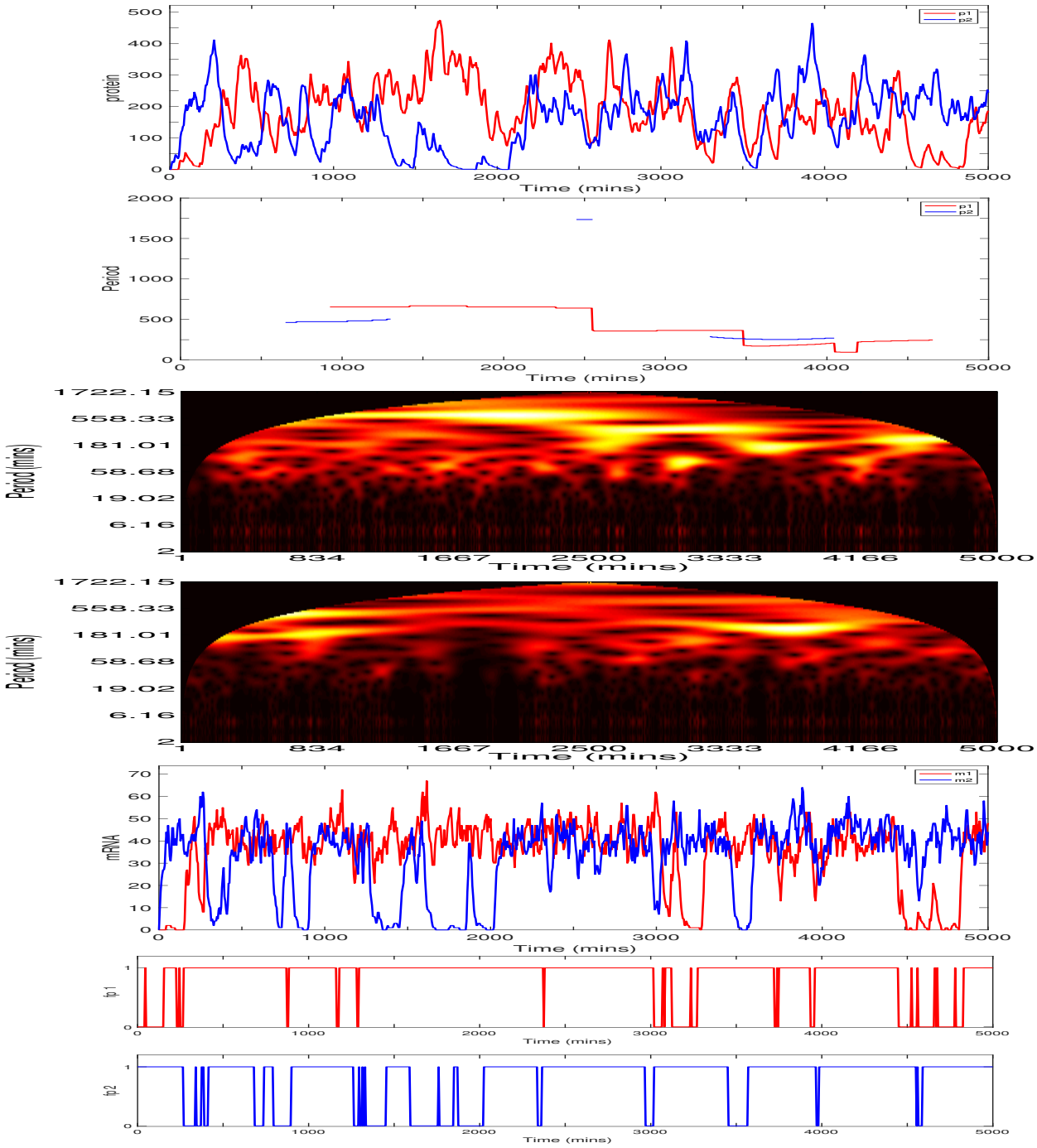

Figure 5.17: *The two gene repressilator with individual gene sites clustered at the centre of the nucleus:* $g_1 = (0.5\mu m, 0\mu m, 0\mu m), g_2 = (-0.5\mu m, 0\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

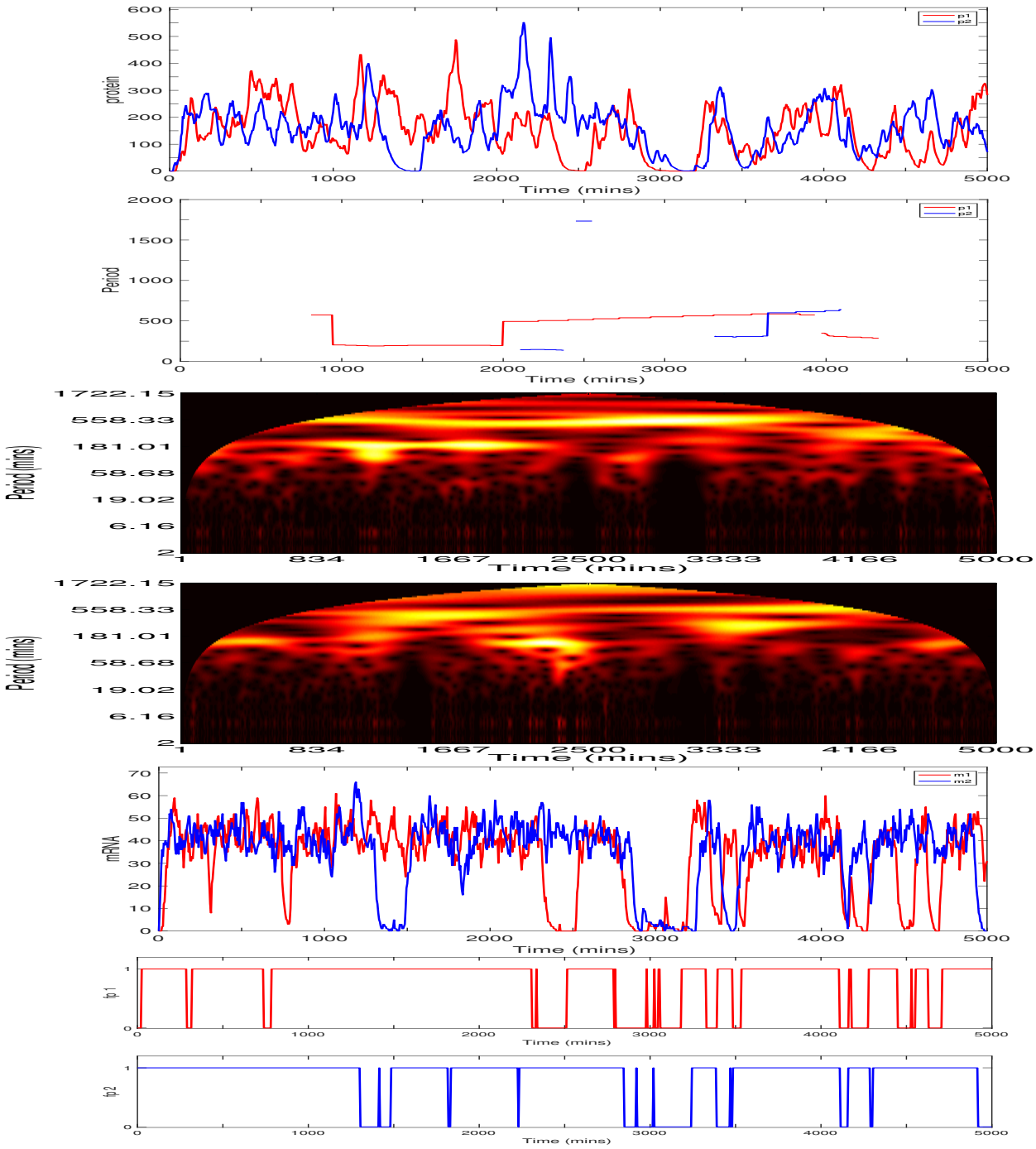

Figure 5.18: *The two gene repressilator with individual gene sites clustered at the centre of the nucleus:* $g_1 = (0.5\mu m, 0\mu m, 0\mu m), g_2 = (-0.5\mu m, 0\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

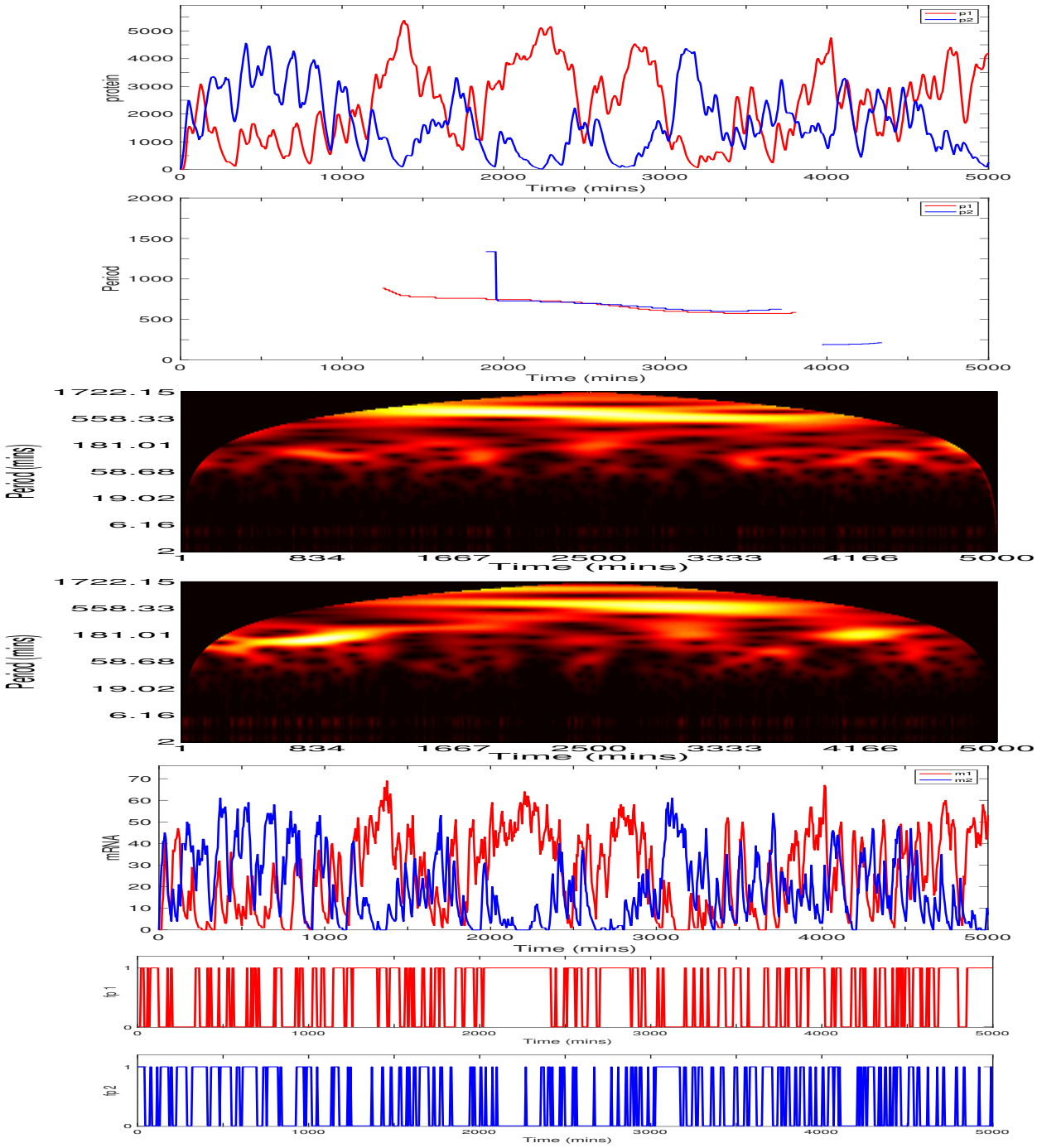

Figure 5.19: *The two gene repressilator with individual gene sites clustered at the centre of the nucleus:* $g_1 = (0.5\mu m, 0\mu m, 0\mu m), g_2 = (-0.5\mu m, 0\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

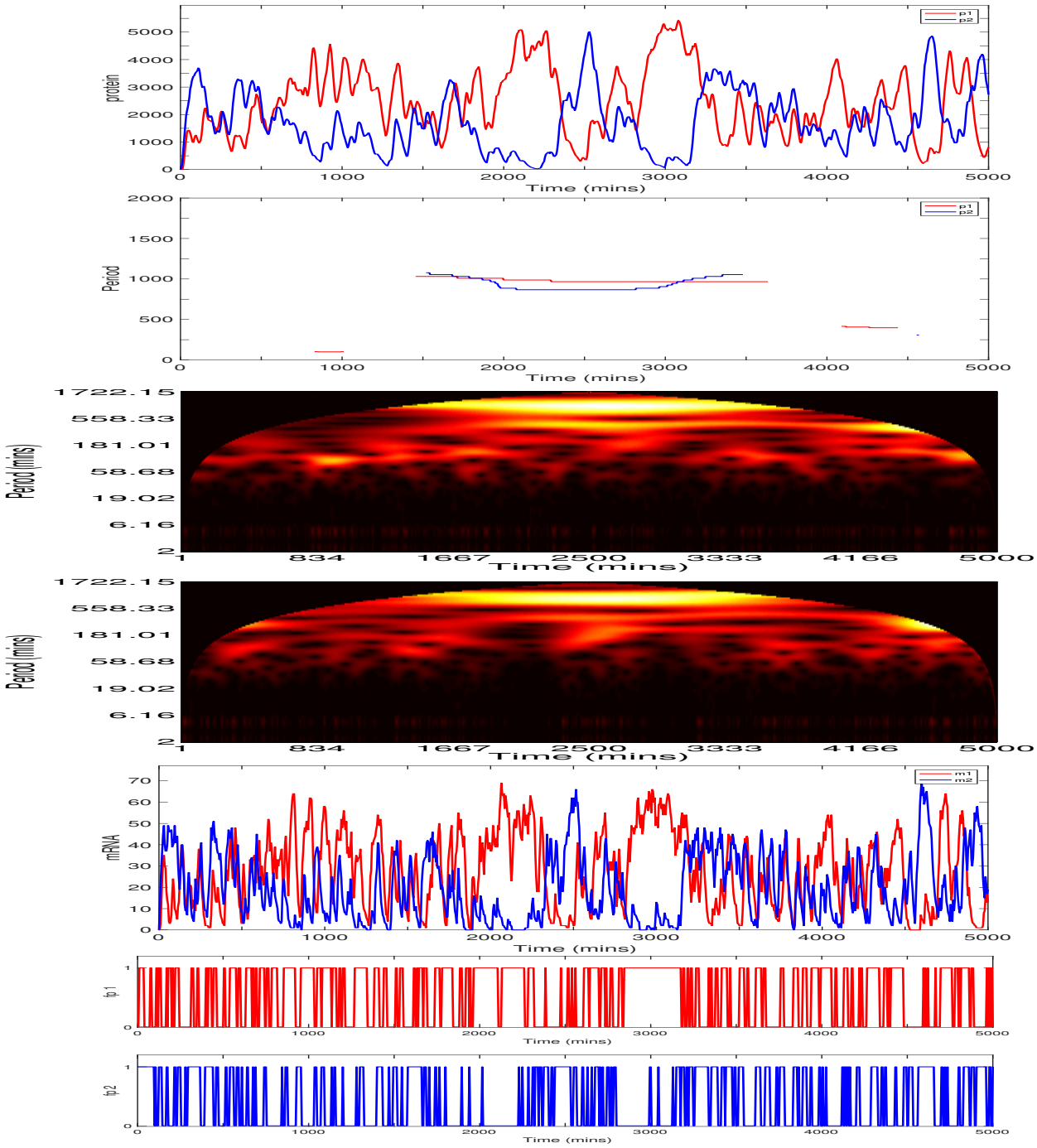

Figure 5.20: *The two gene repressilator with individual gene sites clustered at the centre of the nucleus:* $g_1 = (0.5\mu m, 0\mu m, 0\mu m), g_2 = (-0.5\mu m, 0\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

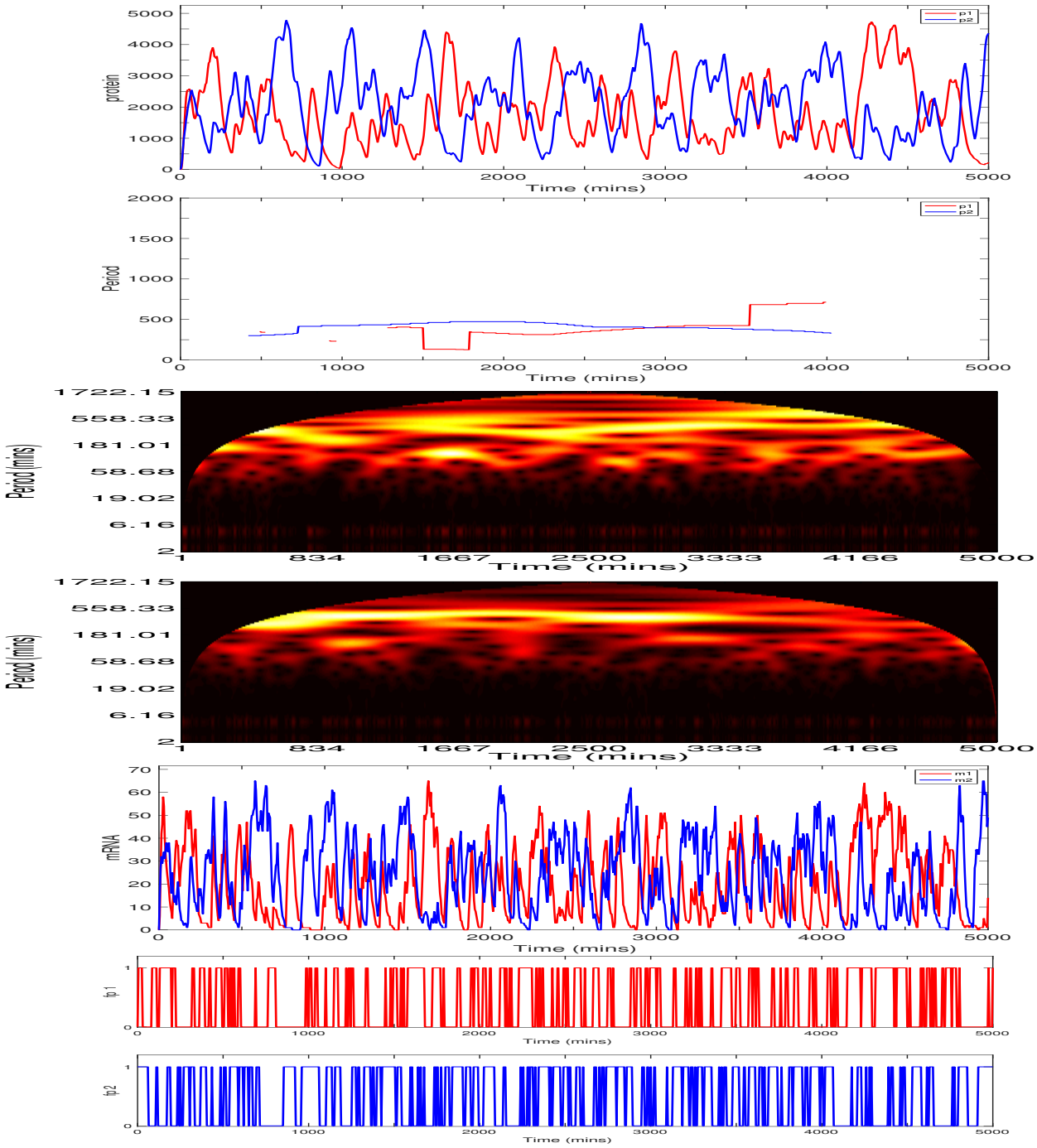

Figure 5.21: *The two gene repressilator with individual gene sites clustered at the centre of the nucleus:* $g_1 = (0.5\mu m, 0\mu m, 0\mu m), g_2 = (-0.5\mu m, 0\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

For $D = 1 \times 10^{-12} m^2 min^{-1}$ the average period for 300 trajectories was $\bar{T} = 504.5$, with their range being, $208.4 \leq \bar{T} \leq 1734$. The average protein number was $\bar{p} = 163.5$, ranging between $116 \leq \bar{p} \leq 232$. The average peak value of protein number was $\bar{p}_{peak} = 455$, ranging between $339 \leq \bar{p}_{peak} \leq 700$.

For $D = 1 \times 10^{-10} m^2 min^{-1}$ the average period for 300 trajectories was $\bar{T} = 636$, ranging between, $203.6 \leq \bar{T} \leq 1734$. The average protein number was $\bar{p} = 1922$, ranging between $1228 \leq \bar{p} \leq 2586$. The average peak value of protein number was $\bar{p}_{peak} = 4764$, ranging between $3526 \leq \bar{p}_{peak} \leq 5823$.

As the value of D is increased by a magnitude of two, the range of periods recorded for both diffusion regimes are similar. However, the average period recorded across trajectories is two hours longer for the faster diffusion regime. for the one-gene repressilator, (gene site at the origin) we found the opposite, that there was a decrease in period for an increase in the diffusion parameter. However, similarly for the one-gene repressilator case, the level of protein numbers increases by a magnitude of one with this increase in diffusion. Once again this is explained by mRNA reaching the cytoplasm faster and hence, for protein synthesis to occur at a higher rate.

Overall, for gene sites clustered near the origin, the behaviour for each species in a two gene repressilator can be compared to the behaviour of a single one-gene repressilator species, with mRNA, protein levels and promoter status behaviours being roughly similar, see figures 5.2, 5.3, 5.4, 5.5, 5.8, 5.9, 5.10 and 5.11. However, please note the different time frame. The one-gene repressilator simulations were ran for 1600 minutes and the two-gene repressilator simulations, (gene sites clustered at the origin) were ran for 5000 minutes. This was in order for WAVOS to detect the higher oscialltory modes that can appear. In consideration of the heat maps, there are a greater number of periods for the two-gene repressilator that are not as well defined than the periods for the one-gene repressilator. However, there appears to be oscillations happening on

a smaller time scale and oscillations that appear on a larger time scale. The range of modal periods is a lot greater in size than the range of modal periods of the one-gene repressilator.

For the faster diffusion regime, both protein species have similar oscillatory behaviour, see figures 5.19, 5.20 and 5.21. One could infer that they oscillate in anti-phase. However, there are long periods of time were one protein has a much larger copy number than the other protein, see figure 5.19 and 5.20. One could infer that the species with lower copy number of protein lies in a lower state and the protein with large copy number lies in the higher state. The presence of the oscillatory behaviour leads to a likelihood that a switching effect can occur. Where the protein with smaller number gains advantage and becomes the larger expressed protein and vice-versa, thus, switching the 'state' of the system. In figures 5.19 and 5.20, protein 2 begins to look dominant but then there is a switch and protein 1 dominates for a long period; a period of time as long as 2000 minutes. This time scale is longer than a day, (1440 minutes). A day, is time scale that most experiments on the toggle switch run for, hence, why the observation of switching effects might be rare. As discussed earlier, Gardner et al. (2000) records that there is bi-stability in the two-gene repressilator but that switching will rarely occur.

However, if we begin to add a factor of spatial-asymmetry, by moving one gene site closer to the nuclear membrane, while fixing the other at the origin, we begin to see how this switching effect becomes less likely, see figures 5.22 and 5.23. Both of these figure are for the smaller diffusion regime $D = 1 \times 10^{-12}$. Figure 5.22 has gene sites at position $g_1 = (0,0,0)$ and $g_2 = (1.5,0,0)$. Figure 5.23 has gene sites at position $g_1 = (0,0,0)$ and $g_2 = (2.85,0,0)$. Proteins now fight for dominance. Once a protein gains advantage, it is highly likely it will remain in the high state. For the latter case, (gene site 2 being close to the nuclear membrane and gene 1 being at the origin), it is almost certain that protein 2 will always take on the higher steady state solution. This

particular finding is in agreement with a lot of previous stochastic models of the toggle switch: that the toggle switch is bi-stable. However, as the diffusion parameter is increased, the somewhat oscillatory behaviour of protein species, (or rather switching effect) is retained.

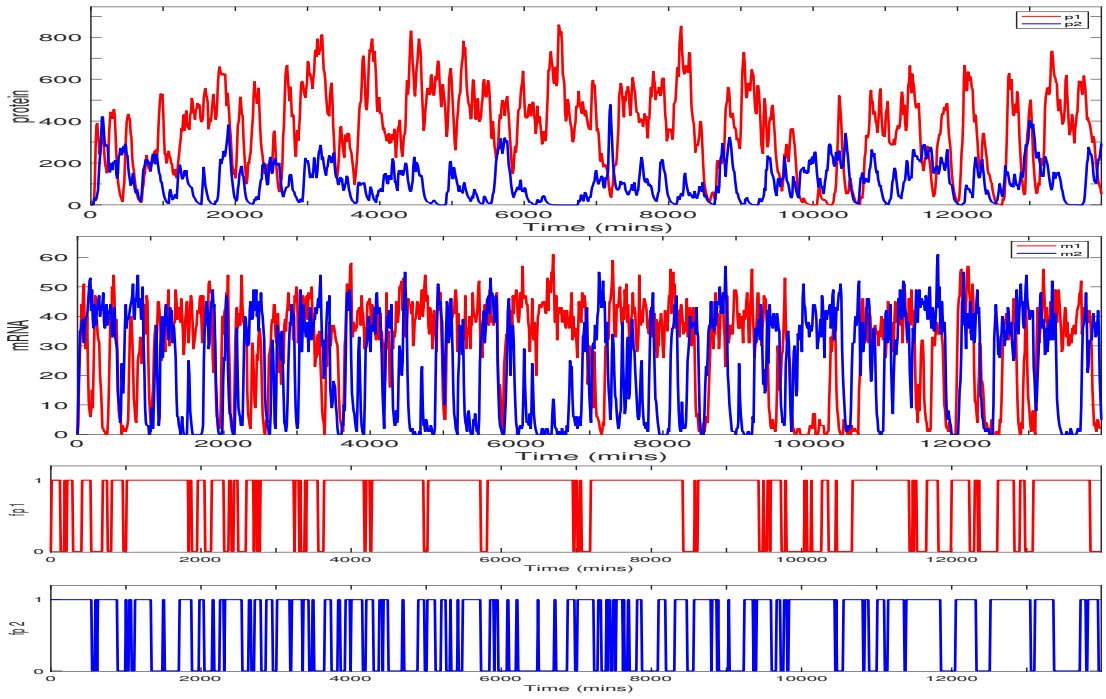

Figure 5.22: *The two-gene repressilator with asymmetric gene site locations:* $g_1 = (0,0,0)$ *and* $g_2 = (1.5,0,0)$ *and diffusion parameter,* $D = 1 \times 10^{-12}$ *Top: protein; middle: mRNA and bottom: free promoter.*

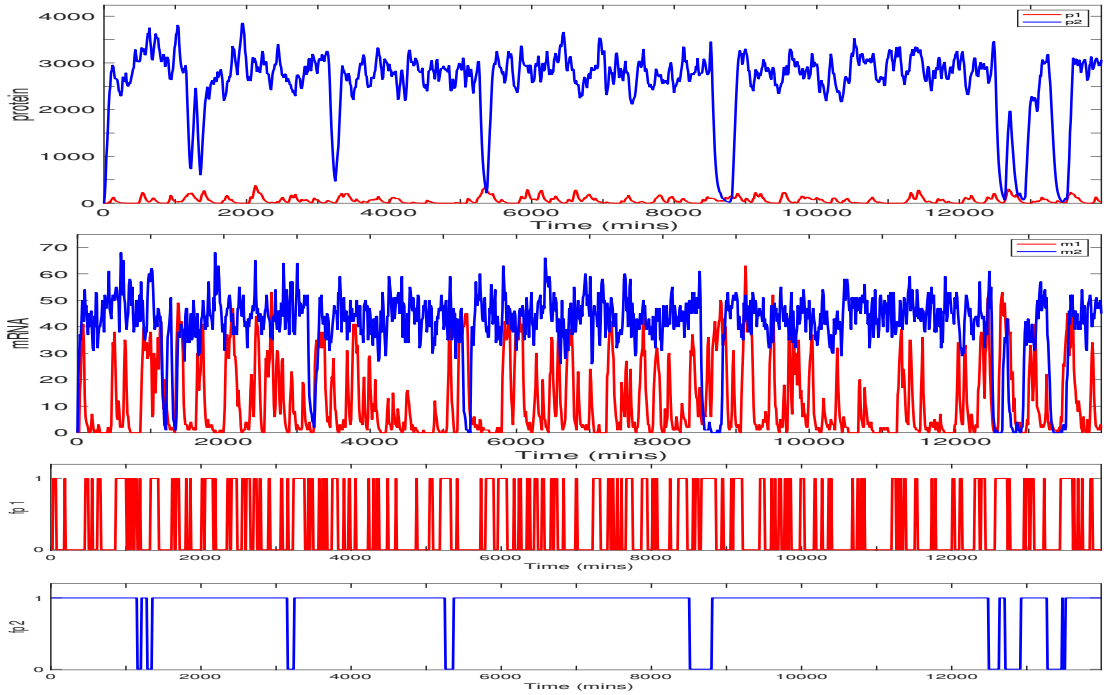

Figure 5.23: *The two-gene repressilator with asymmetric gene site locations:* $g_1 = (0,0,0)$ *and* $g_2 = (2.85,0,0)$ *and diffusion parameter,* $D = 1 \times 10^{-12}$ *Top: protein; middle: mRNA and bottom: free promoter.*

We now consider the two-gene repressilator with gene site locations at $g_1 = (x_1, y_1, z_3) = (2.85\mu m, 0\mu m, 0\mu m)$ and $g_2 = (x_2, y_2, z_3) = (-2.85\mu m, 0\mu m, 0\mu m)$ for gene 1 and gene 2 respectively. Hence, we move them quite far apart from each other, each close to the nuclear membrane. The first set of three simulations are for the smaller diffusion regime, $D = 1 \times 10^{-12}$ and the second set of three simulations are for the faster diffusion regime, $D = 1 \times 10^{-10}$.

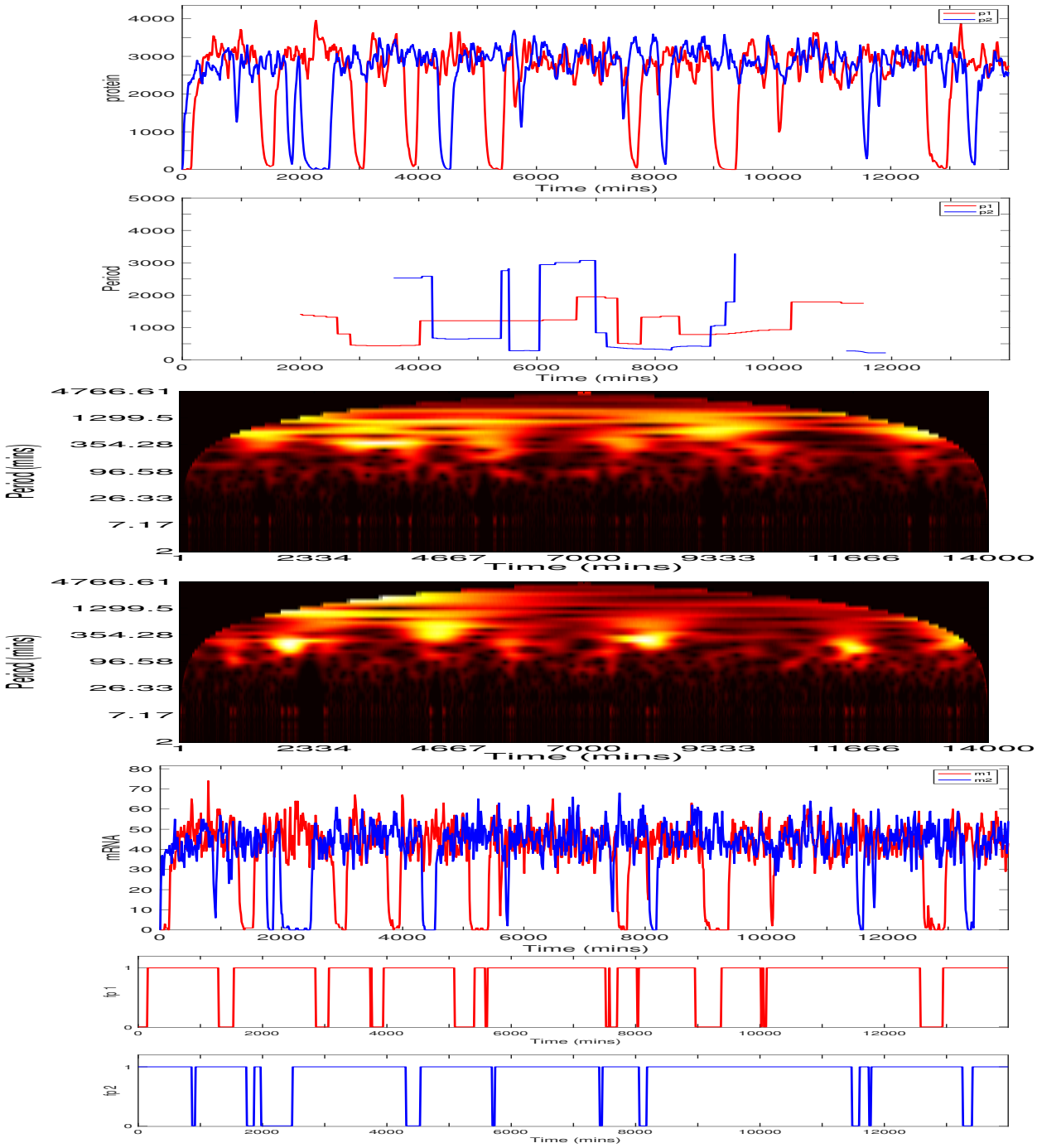

Figure 5.24: *The two gene repressilator with individual gene sites equidistant apart from the origin:* $g_1 = (2.85\mu m, 0\mu m, 0\mu m), g_2 = (-2.85\mu m, 0\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

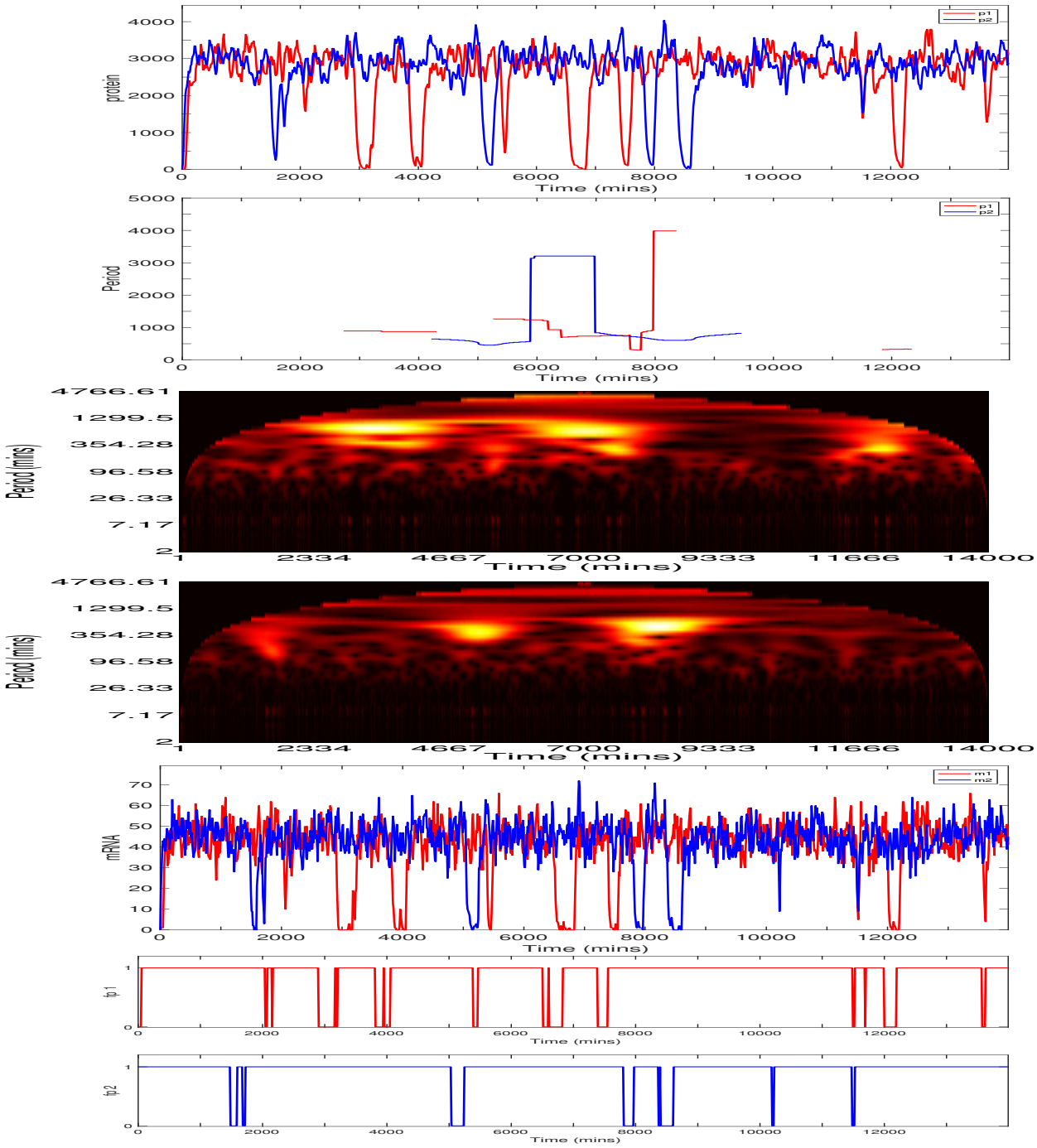

Figure 5.25: *The two gene repressilator with individual gene sites equidistant apart from the origin:* $g_1 = (2.85\mu m, 0\mu m, 0\mu m), g_2 = (-2.85\mu m, 0\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

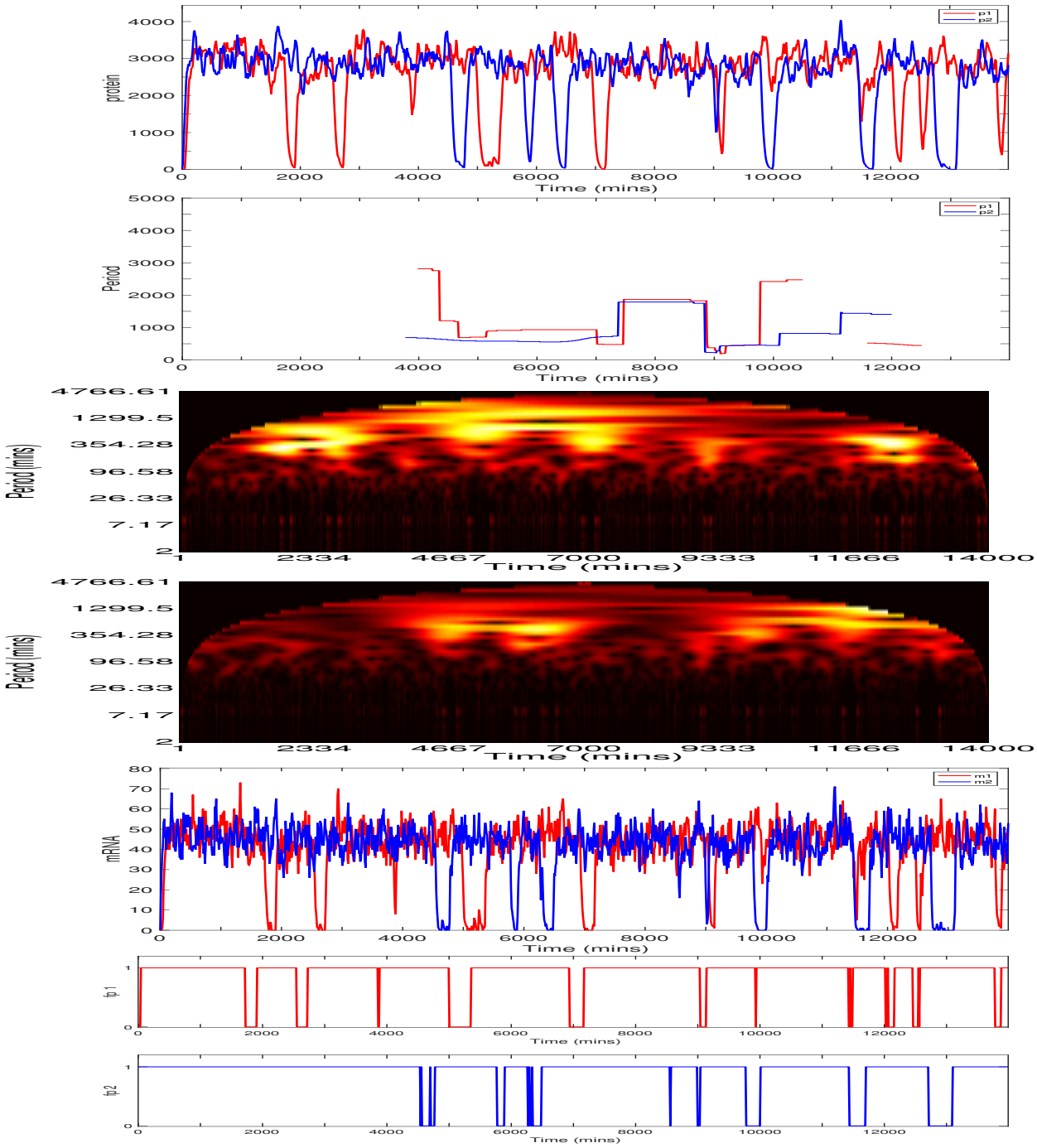

Figure 5.26: *The two gene repressilator with individual gene sites equidistant apart from the origin:* $g_1 = (2.85\mu m, 0\mu m, 0\mu m), g_2 = (-2.85\mu m, 0\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

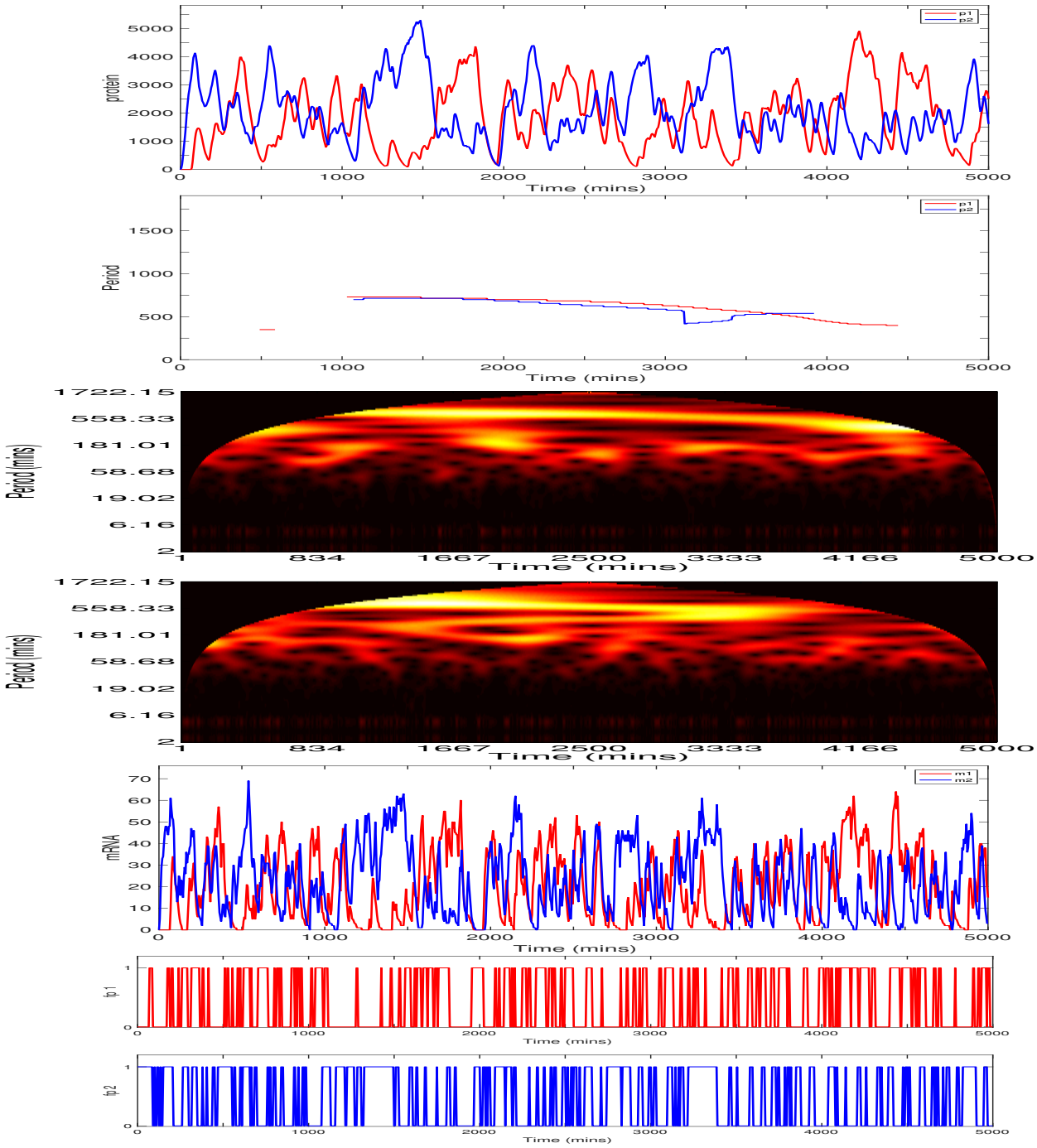

Figure 5.27: *The two gene repressilator with individual gene sites equidistant apart from the origin:* $g_1 = (2.85\mu m, 0\mu m, 0\mu m), g_2 = (-2.85\mu m, 0\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

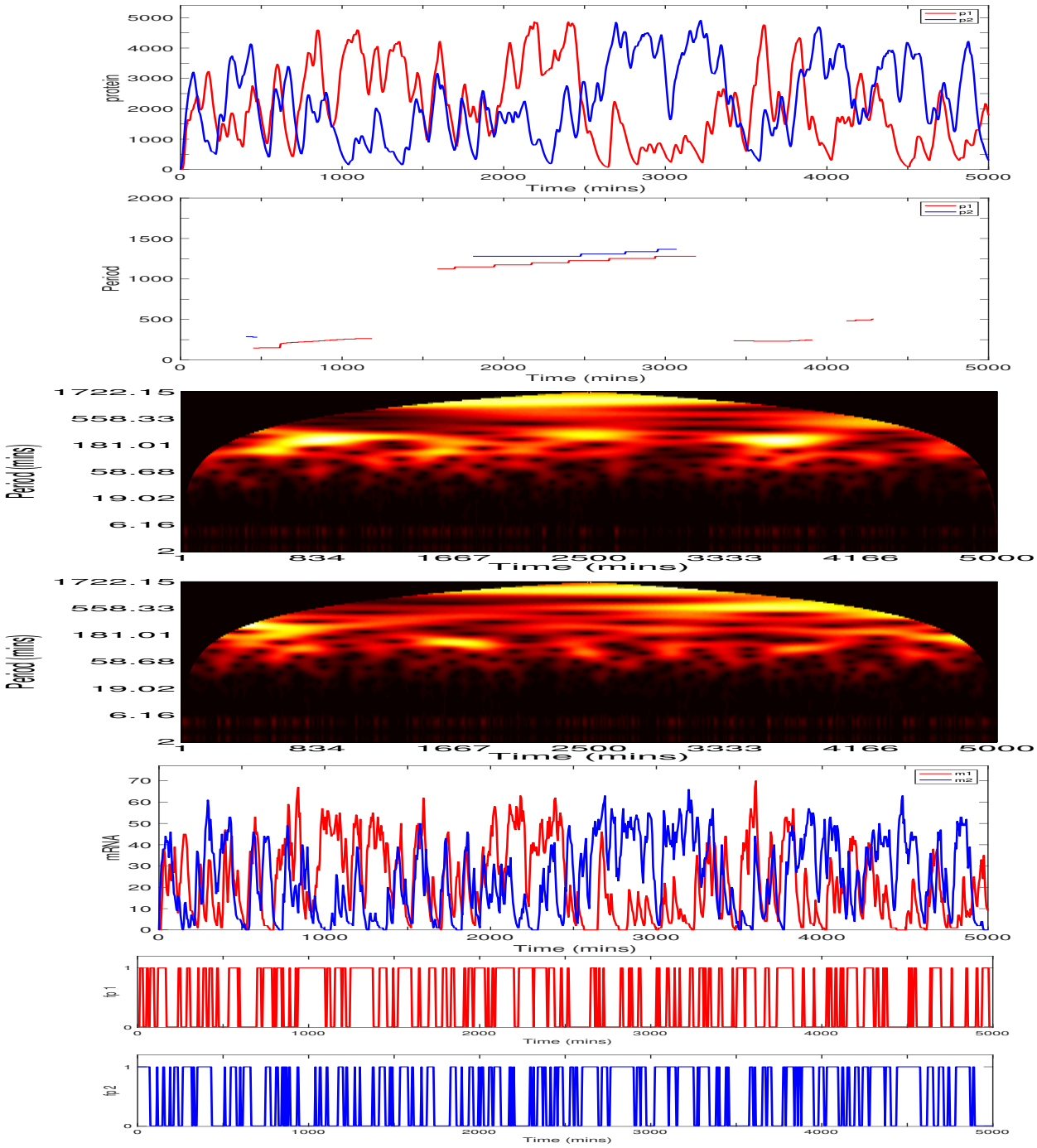

Figure 5.28: *The two gene repressilator with individual gene sites equidistant apart from the origin:* $g_1 = (2.85\mu m, 0\mu m, 0\mu m), g_2 = (-2.85\mu m, 0\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

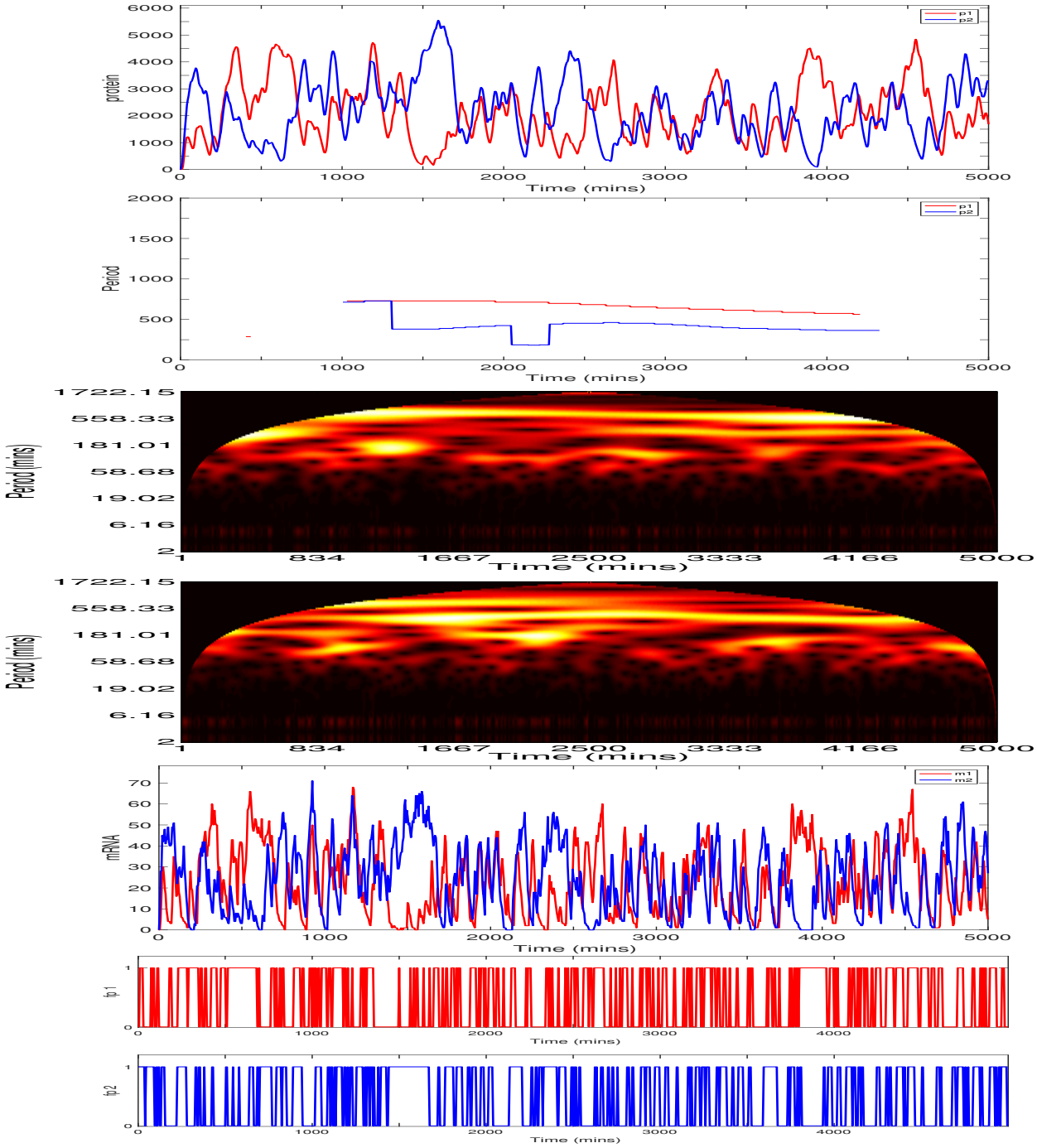

Figure 5.29: *The two gene repressilator with individual gene sites equidistant apart from the origin:* $g_1 = (2.85\mu m, 0\mu m, 0\mu m), g_2 = (-2.85\mu m, 0\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

For $D = 1 \times 10^{-12} m^2 min^{-1}$ the average period for 200 trajectories was $\bar{T} = 1128$, ranging between, $300 \leq \bar{T} \leq 4015$. The average protein number was $\bar{p} = 2625$, ranging between $2241 \leq \bar{p} \leq 2826$. The average peak value of protein number was $\bar{p}_{peak} = 3886$, ranging between $3572 \leq \bar{p}_{peak} \leq 4375$.

For $D = 1 \times 10^{-10} m^2 min^{-1}$ the average period for 200 trajectories was $\bar{T} = 595$, ranging between, $205 \leq \bar{T} \leq 1618$. The average protein number was $\bar{p} = 2015$, ranging between $1388 \leq \bar{p} \leq 2604$. The average peak value of protein number was $\bar{p}_{peak} = 5005$, ranging between $4161 \leq \bar{p}_{peak} \leq 5927$.

For the faster diffusion regime, we can compare these trajectories, (gene sites close to the membrane, equidistant along the positive and negative x-axis) to the simulations with gene sites clustered at the origin, there is relatively no difference. However, the former has a slightly faster period, suggesting it is a better 'oscillator', although still a weak one. For the smaller diffusion regime, we see quite different behaviour. Both protein species tend to stay quite highly expressed. This is due to the fact that each of there promoters is located next to the nuclear membrane such that protein can be maximally expressed. However, because the target promoter is far across the nucleus, there is less protein locating in that area, due to the smaller diffusion parameter. Hence, the lack of negative feedback. It is quite a stretch to call this scenario an oscillator. Although, protein species do get turned off for a short while at what appears to be random times.

We now consider the gene site locations to be $g_1 = (x_1, y_1, z_1) = (0.5\mu m, 2.85\mu m, 0\mu m)$ and $g_2 = (x_2, y_2, z_3) = (-0.5\mu m, 2.85\mu m, 0\mu m)$ for gene 1 and gene 2 respectively i.e. clustered close to the nuclear membrane. The first set of three simulations are for the smaller diffusion regime, $D = 1 \times 10^{-12}$ and the second set of three simulations are for the faster diffusion regime, $D = 1 \times 10^{-10}$.

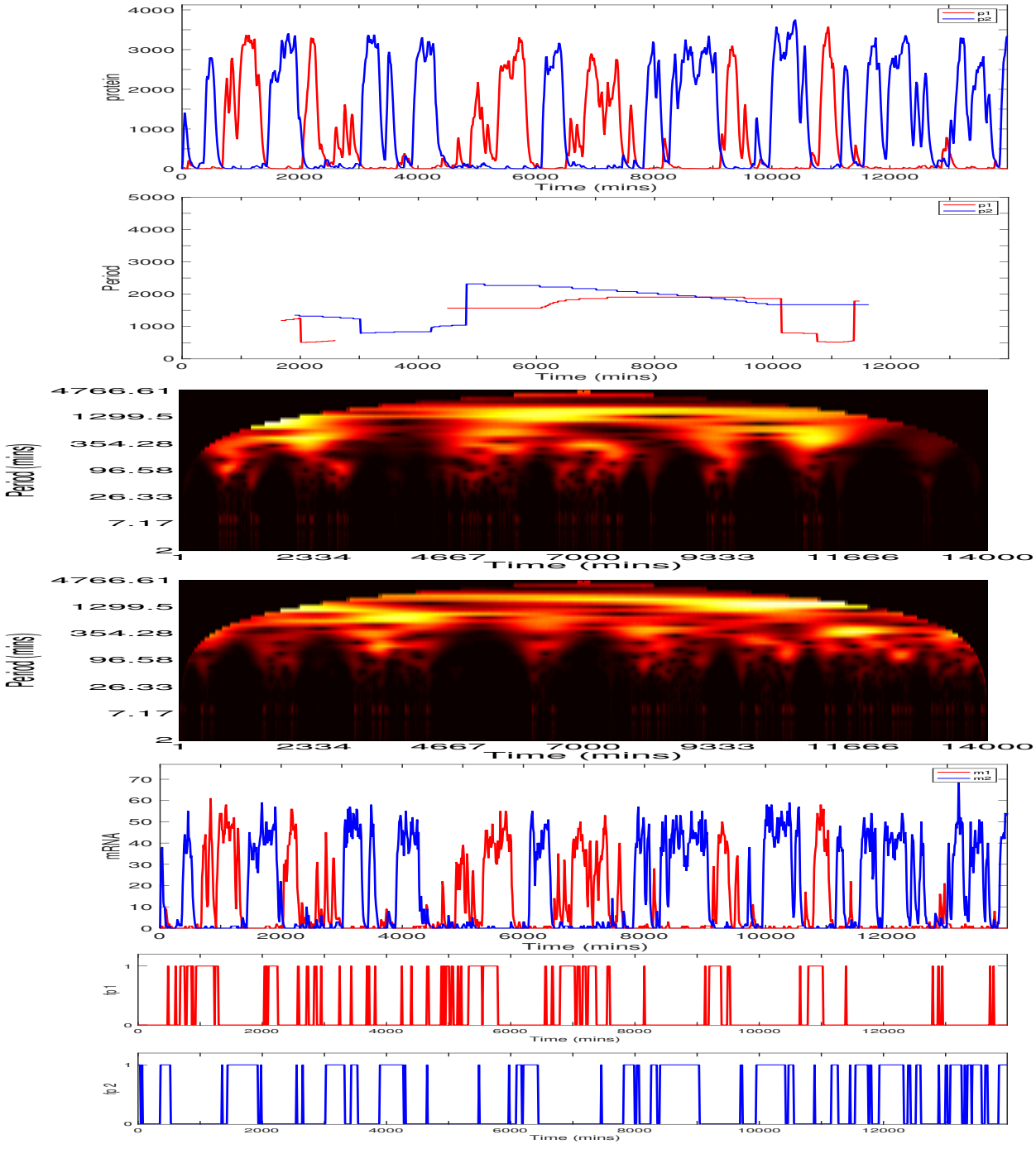

Figure 5.30: *The two gene repressilator with individual gene sites:* $g_1 = (0.5\mu m, 2.85\mu m, 0\mu m), g_2 = (-0.5\mu m, 2.85\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: Protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

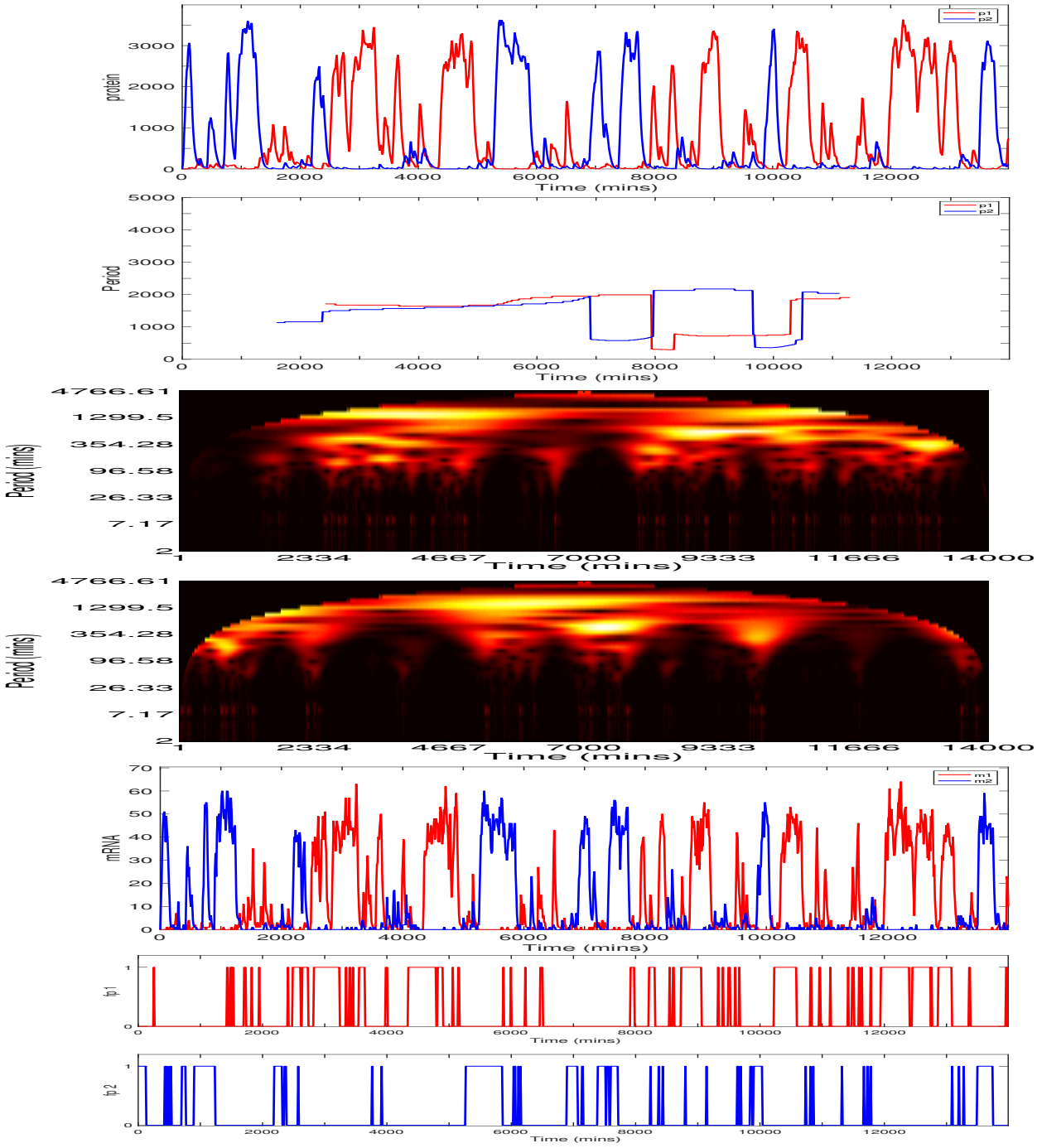

Figure 5.31: *The two gene repressilator with individual gene sites:* $g_1 = (0.5\mu m, 2.85\mu m, 0\mu m), g_2 = (-0.5\mu m, 2.85\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: Protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

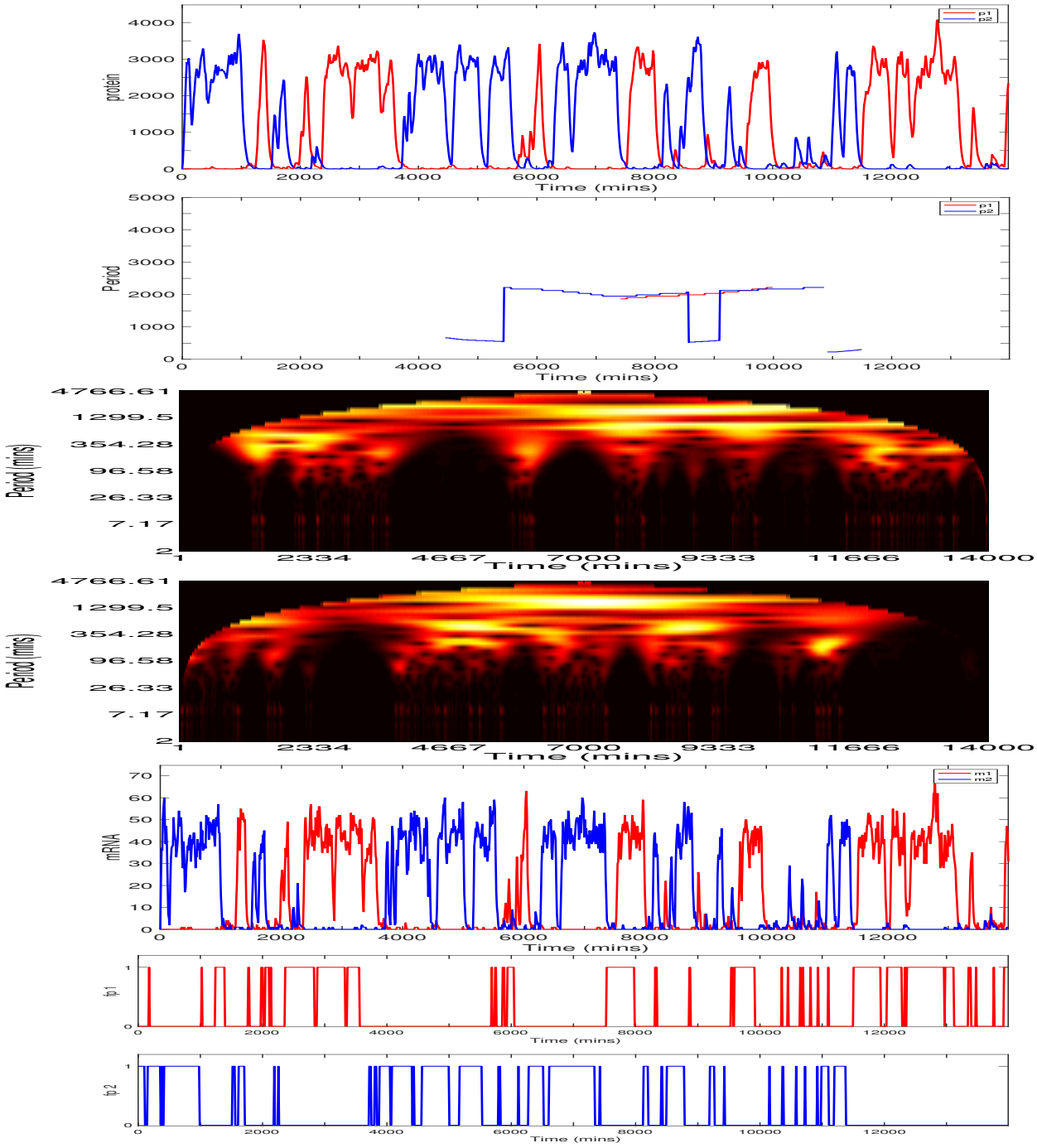

Figure 5.32: *The two gene repressilator with individual gene sites:* $g_1 = (0.5\mu m, 2.85\mu m, 0\mu m), g_2 = (-0.5\mu m, 2.85\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: Protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

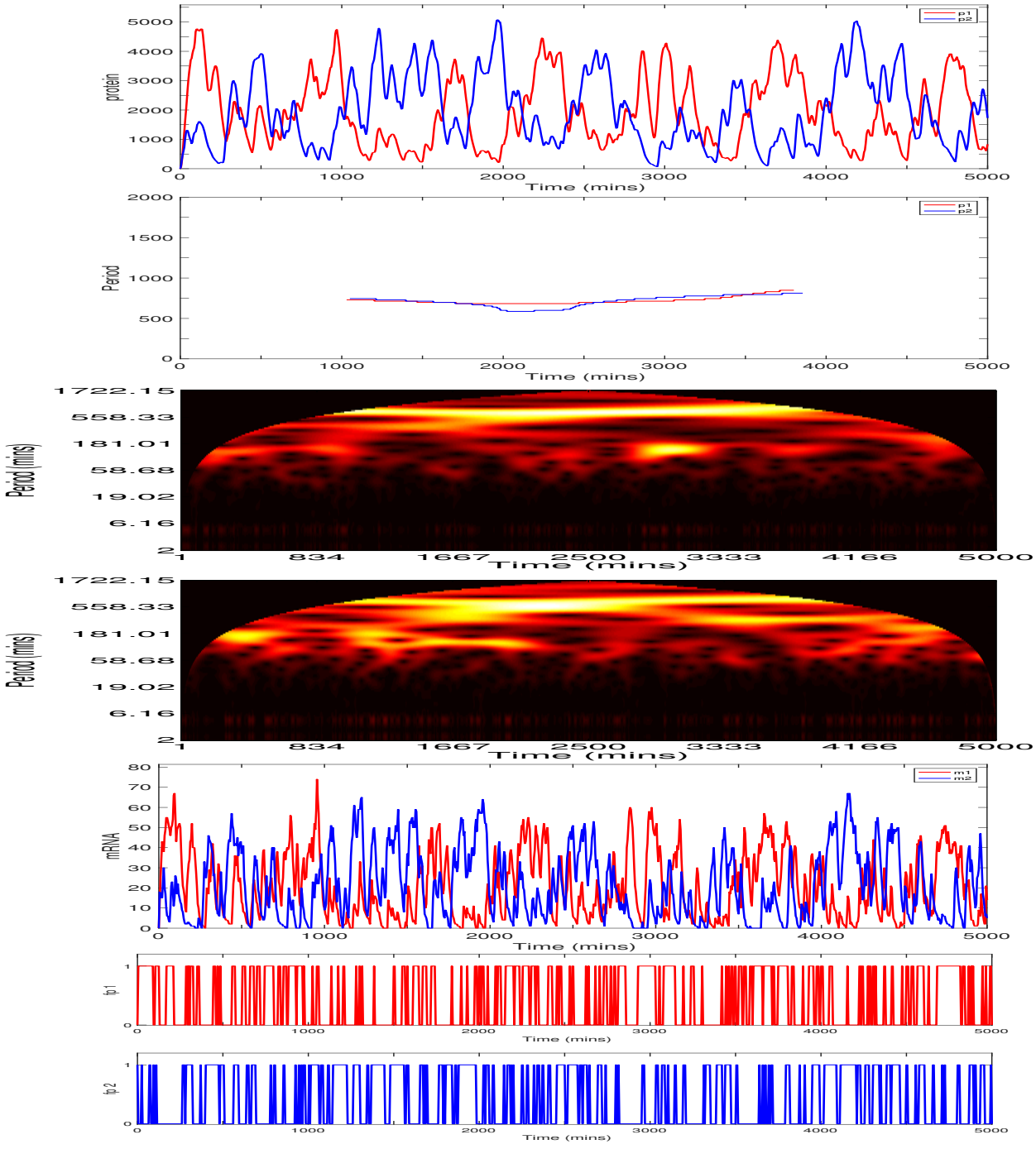

Figure 5.33: *The two gene repressilator with individual gene sites:* $g_1 = (0.5\mu m, 2.85\mu m, 0\mu m), g_2 = (-0.5\mu m, 2.85\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: Protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

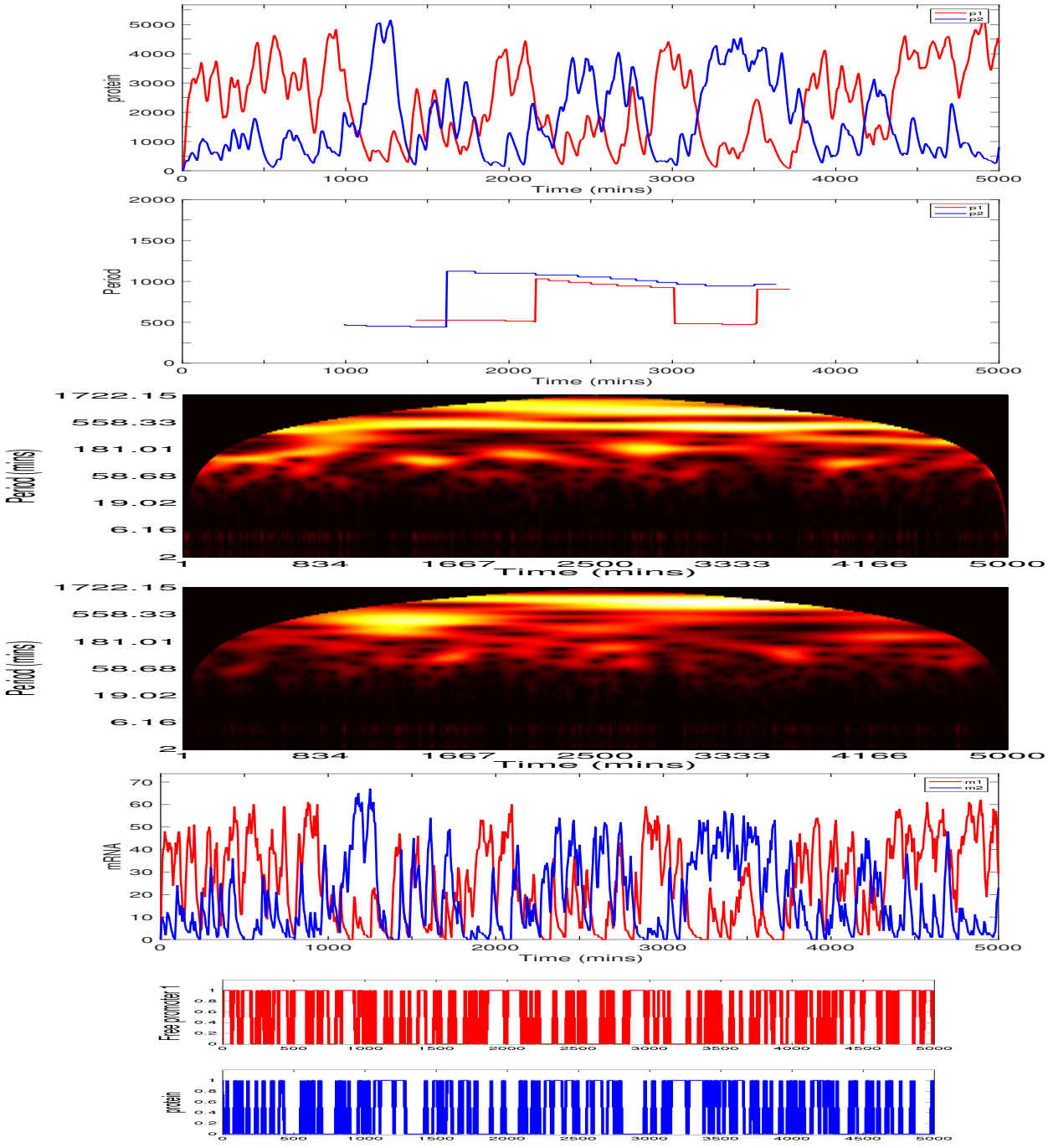

Figure 5.34: *The two gene repressilator with individual gene sites:* $g_1 = (0.5\mu m, 2.85\mu m, 0\mu m), g_2 = (-0.5\mu m, 2.85\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: Protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

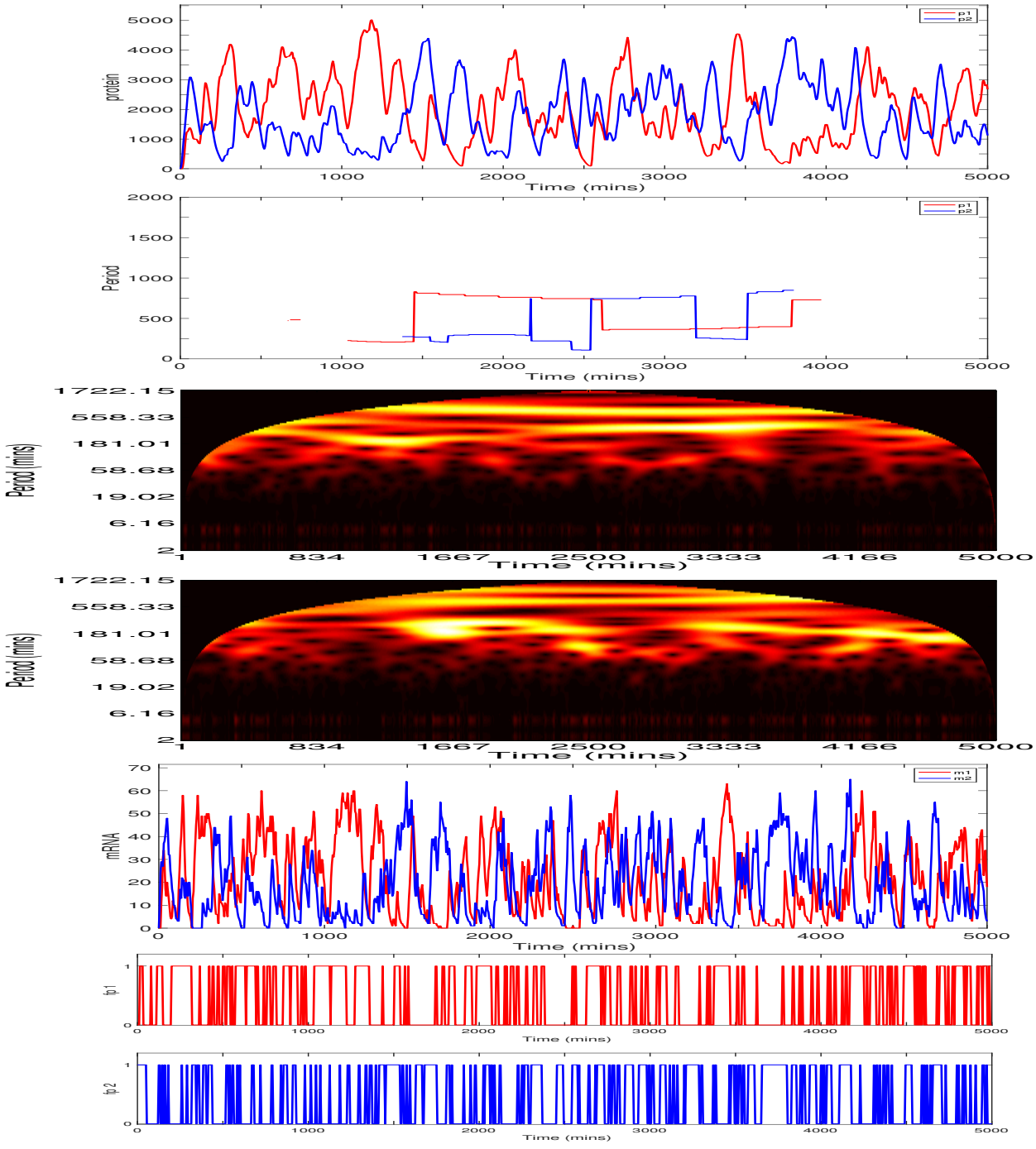

Figure 5.35: *The two gene repressilator with individual gene sites:* $g_1 = (0.5\mu m, 2.85\mu m, 0\mu m), g_2 = (-0.5\mu m, 2.85\mu m, 0\mu m)$ *and a diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, and Gene2/products. Top to bottom: Protein, period, heat map of protein period 1 and 2, mRNA and free promoter time series. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

For $D = 1 \times 10^{-12} m^2 min^{-1}$ the average period for 200 trajectories was $\bar{T} = 1553$, with their range being, $434 \leq \bar{T} \leq 4161$. The average protein number was $\bar{p} = 829$, ranging between $142 \leq \bar{p} \leq 1730$. The average peak value of protein number was $\bar{p}_{peak} = 3612$, ranging between $2635 \leq \bar{p}_{peak} \leq 4261$.

For $D = 1 \times 10^{-10} m^2 min^{-1}$ the average period for 200 trajectories was $\bar{T} = 661.4$, ranging between, $224 \leq \bar{T} \leq 1734$. The average protein number was $\bar{p} = 1899$, ranging between $1354 \leq \bar{p} \leq 2555$. The average peak value of protein number was $\bar{p}_{peak} = 4962$, ranging between $4170 \leq \bar{p}_{peak} \leq 5860$.

As D is increased from $1 \times 10^{-12} m^2 min^{-1}$ to $1 \times 10^{-10} m^2 min^{-1}$ the average instantaneous period decreases quite drastically, (around 15 hours).

For both diffusion regimes, protein for each species can be highly expressed. However, there is also a slight increase in protein levels as D is increased. This again is due to the fact that more mRNA reach the cytoplasm and hence protein to be synthesised at a greater rate. The fact that the protein level is high, there is a higher level of fluctuation of protein-promoter binding for both regimes. Both appear to oscillate, (or rather switch between high and low levels of protein). Simulations in the higher diffusion regime show that the individual protein species oscillate in their levels at a higher rate than in the slower diffusion regime. This indicates that for the higher diffusion regime, switching will happen with a higher probability.

For the two gene repressilator we suggest that the addition of diffusion into our model stimulates stochastic switching. As the diffusion parameter D is increased, the likelihood of switching between 'states' increases. An increase in diffusion is equivalent to moving the gene sites, as a cluster, closer to the nuclear membrane. Hence, moving gene sites closer to the nuclear membrane should also stimulate stochastic switching more. However, for the smaller diffusion regime, there were situations in which

switch-like behaviour would be extremely rare. This was either for the gene sites being at different distances away from the cytoplasm, such that one protein could gain a dominant expression level and hence, dominantly repress the other. Or when both gene sites were close to the nuclear membrane but maximally apart from each other along the x-axis. For the first, only one protein was in the higher state and the other in the lower. For the second, both gene sites were close to the nuclear membrane but however, too far away from the other target gene. Thus, both expressed happily in the absence of the other.

However, the findings are supported by Tian & Burrage (2006), Wang et al. (2007) and Xu et al. (2014). They stress that due to noise stochastic switching may occur. Although, the time of switching may be long. We found that the time of switching decreased for an increase in the diffusion parameter.

However, it would be highly unlikely that a cell would have the structure such that gene sites are located with symmetry. It is also highly unlikely, given the crowding of the cell such that protein could diffuse at such a high rate. The rate of diffusion and the position of the gene site needed to be optimal for switch-like behaviour to occur in this model.

Figure 5.36 shows a summary of the mean protein expression levels and mean periods, across 200 trajectories, against both diffusion regimes. Red indicates gene sites clustered at the origin, green indicates gene sites clustered at the nuclear membrane and blue indicates spread apart along the x-axis at the nuclear membrane. We generally find that there is an increase in protein levels as the diffusion parameters increased. Also there is an increase in protein levels for a decrease in the distance of the gene sites to the cytoplasm. However, this is not so much the case when both gene sites were in a cluster near the nuclear membrane, as negative feedback on each gene is likely to happen at a faster rate, due to the closer proximity. For the faster diffusion regime

there was relatively no change in oscillatory (or switch-like) behaviour, however, for the slower diffusion regime, the likelihood of switching decreases as the gene sites are moved closer to the nuclear membrane.

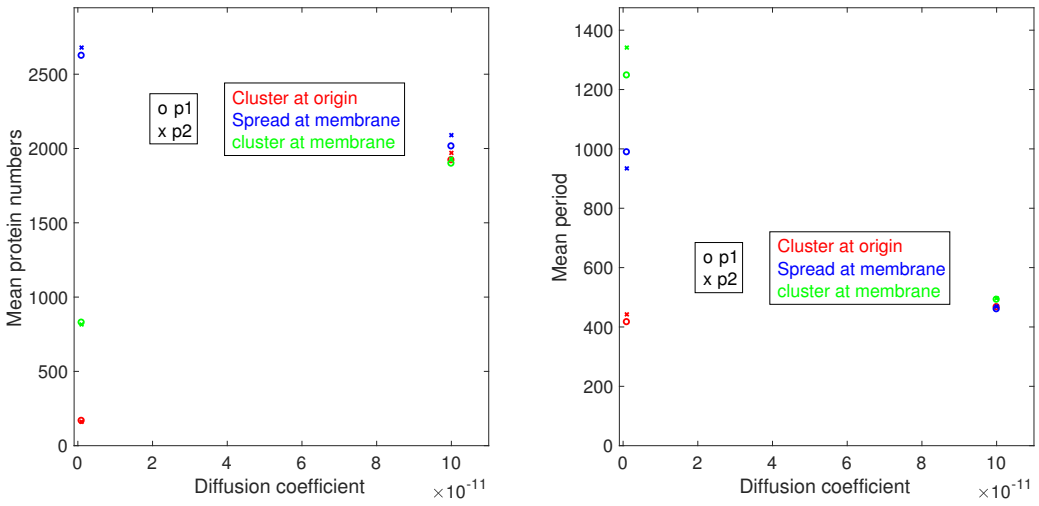

Figure 5.36: *Plots to summarise the two-gene repressilator. Left: Plot of mean protein levels agains diffusion. Right: Plot of mean period against diffusion. Red indicates gene sites clustered at the origin, green indicates gene sites clustered at the nuclear membrane and blue indicates spread apart along the x-axis at the nuclear membrane.*

5.3 The three gene repressilator

We now present the spatial-stochastic model of the three gene repressilator. We first consider the gene site locations to be $g_1 = (x_1, y_1, z_1) = (0\mu m, \frac{1}{\sqrt{3}}\mu m, 0\mu m)$, $g_2 = (x_2, y_2, z_3) = (0.5\mu m, -\frac{1}{2\sqrt{3}}\mu m, 0\mu m)$ and $g_3 = (x_3, y_3, z_3) = (-0.5\mu m, -\frac{1}{2\sqrt{3}})$ for gene 1, 2 and 3 respectively. Thus, each gene sites lies at a distance equidistant away from the origin, on the vertices of a triangle with sides of length $1\mu m$.

The first set of three simulations are for the smaller diffusion regime, $D = 1 \times 10^{-12}$ and the second set of three simulations are for the faster diffusion regime, $D = 1 \times 10^{-10}$. We run simulation under the slow diffusion regime for 5000 minutes and simulations under the faster diffusion regime for 4000 minutes. We suspect that the smaller diffusion regime will have higher modal periods that we will want WAVOS to detect.

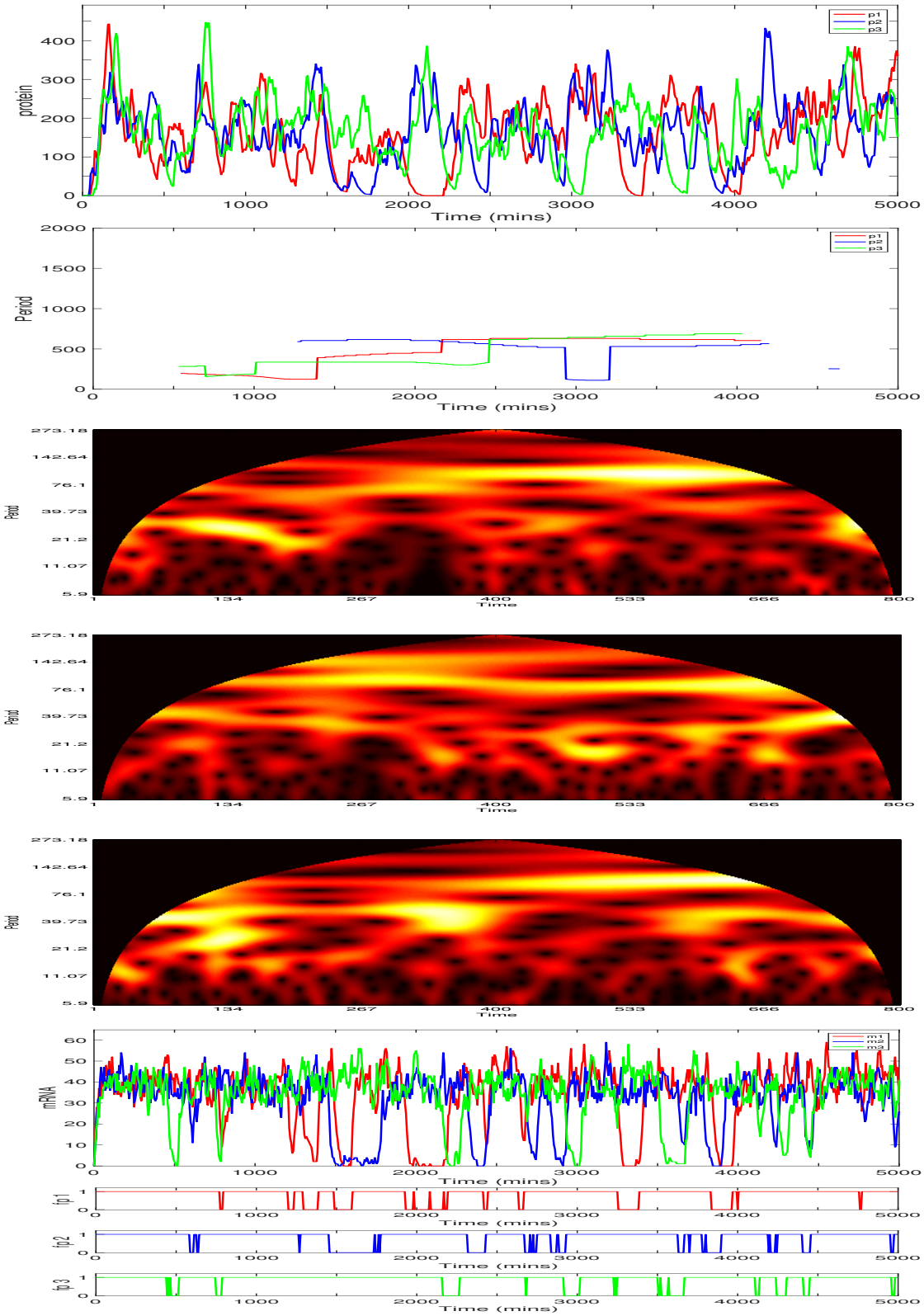

Figure 5.37: *The three gene repressilator with individual gene sites clustered together at the origin:* $g_1 = (0\mu m, 0.5\mu m, 0\mu m), g_2 = (0.5\cos(\pi/6)\mu m, -0.5\sin(\pi/6)\mu m, 0\mu m)$ *and* $g_3 = (-0.5\cos(\pi/6)\mu m, -0.5\sin(\pi/6)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 6.25 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 6.25. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

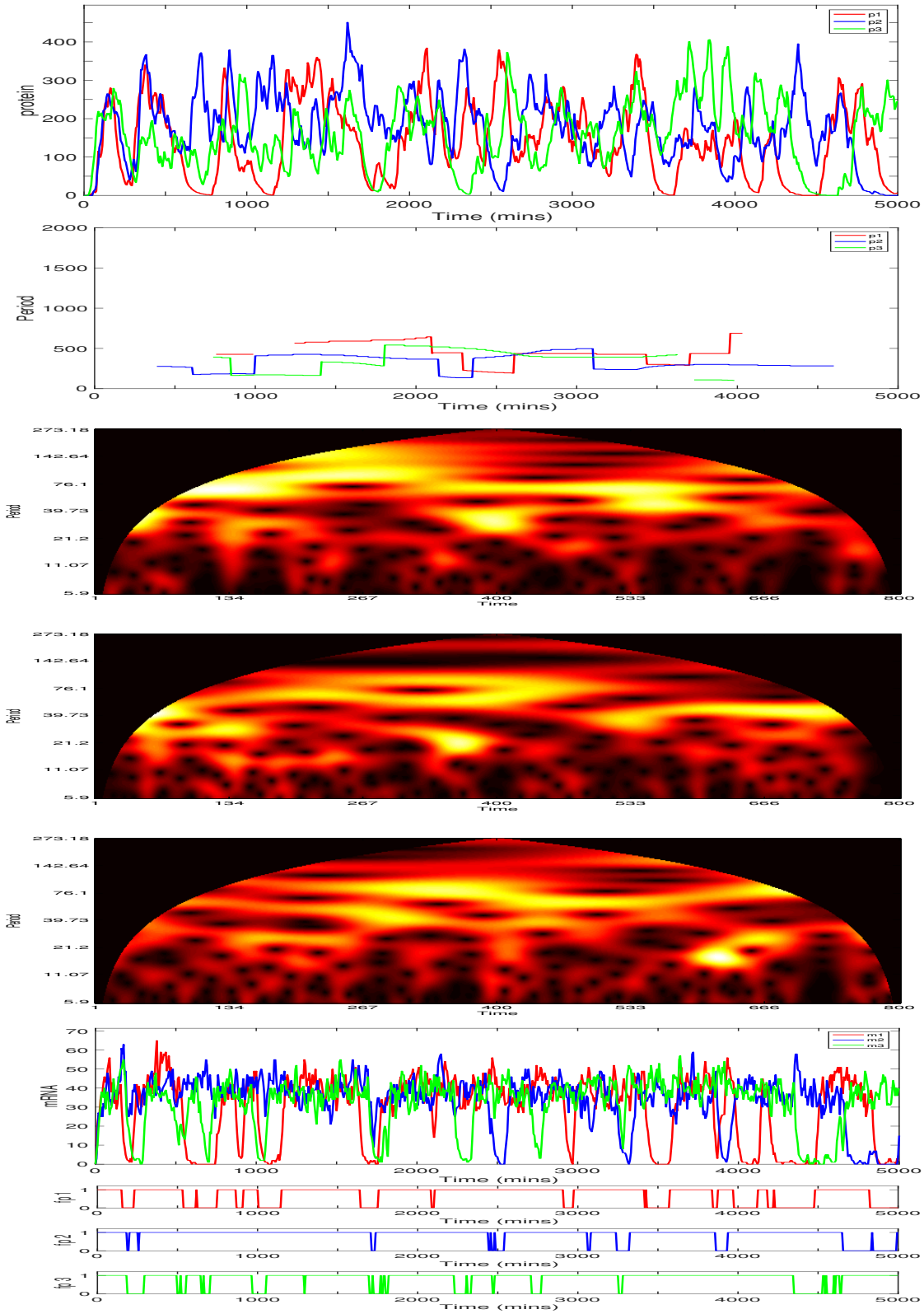

Figure 5.38: *The three gene repressilator with individual gene sites clustered together at the origin:* $g_1 = (0\mu m, 0.5\mu m, 0\mu m), g_2 = (0.5\cos(\pi/6)\mu m, -0.5\sin(\pi/6)\mu m, 0\mu m)$ *and* $g_3 = (-0.5\cos(\pi/6)\mu m, -0.5\sin(\pi/6)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 6.25 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 6.25. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

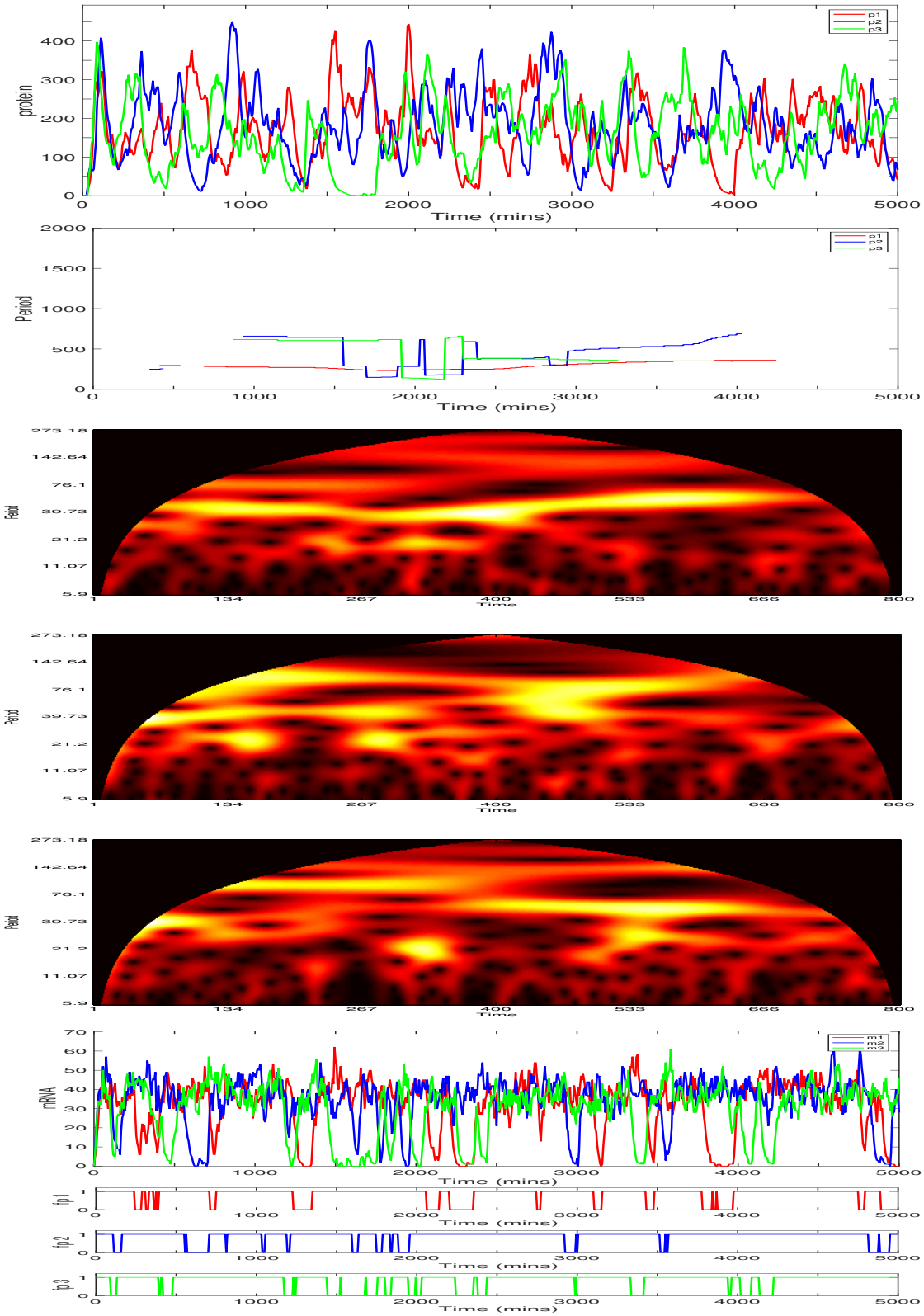

Figure 5.39: *The three gene repressilator with individual gene sites clustered together at the origin:* $g_1 = (0\mu m, 0.5\mu m, 0\mu m), g_2 = (0.5\cos(\pi/6)\mu m, -0.5\sin(\pi/6)\mu m, 0\mu m)$ *and* $g_3 = (-0.5\cos(\pi/6)\mu m, -0.5\sin(\pi/6)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 6.25 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 6.25. allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

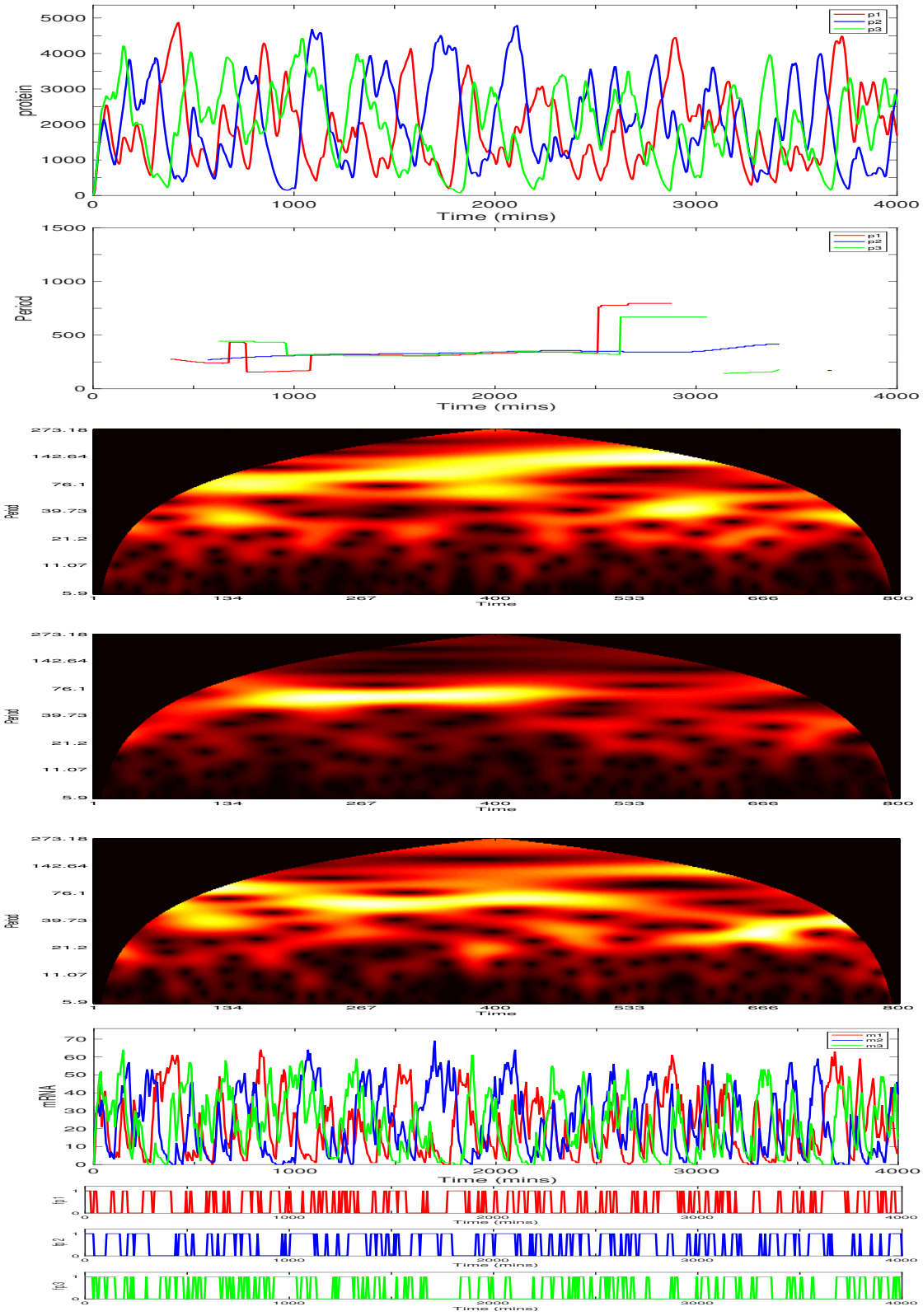

Figure 5.40: *The three gene repressilator with individual gene sites clustered together at the origin:* $g_1 = (0\mu m, 0.5\mu m, 0\mu m), g_2 = (0.5\cos(\pi/6)\mu m, -0.5\sin(\pi/6)\mu m, 0\mu m)$ *and* $g_3 = (-0.5\cos(\pi/6)\mu m, -0.5\sin(\pi/6)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

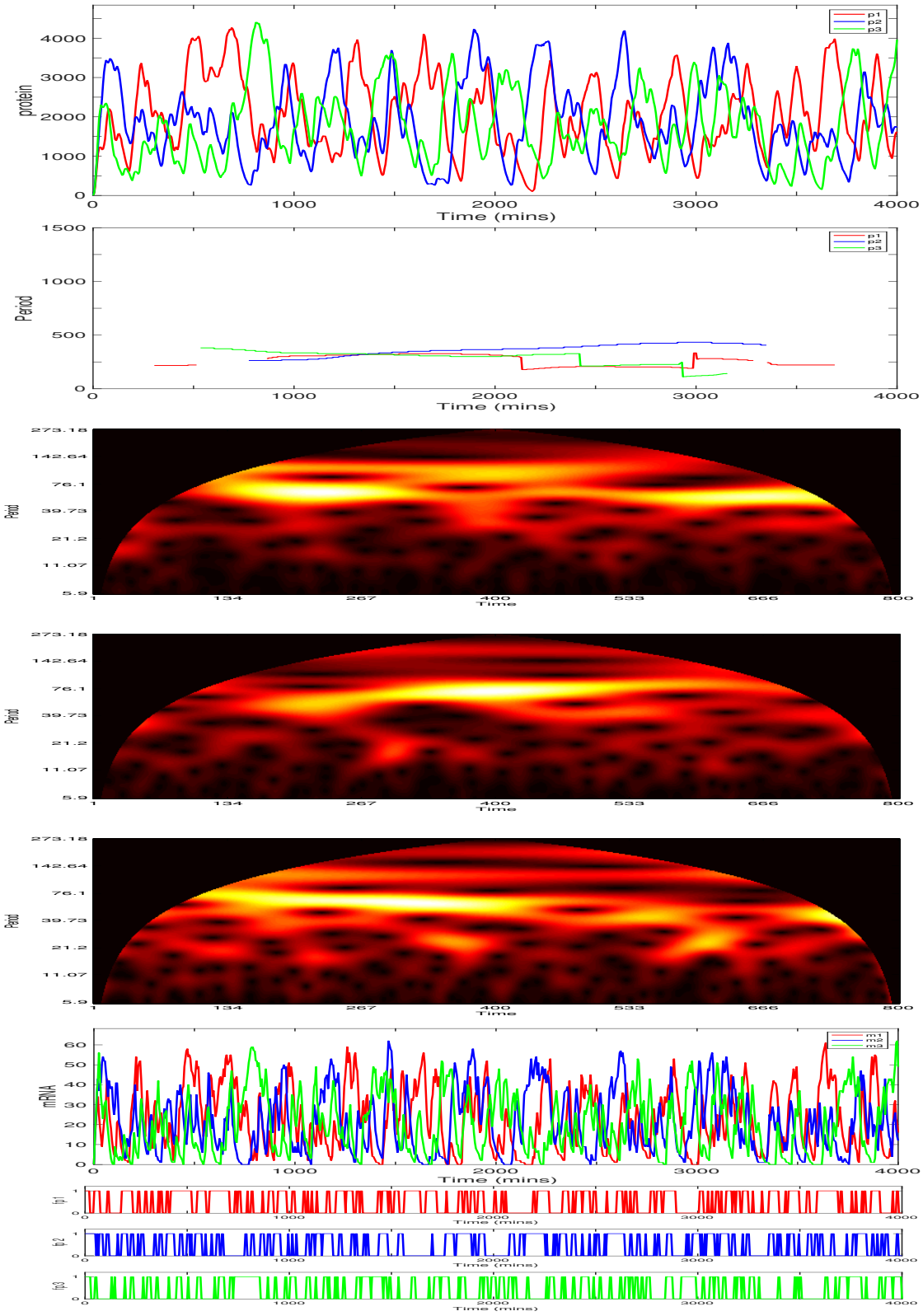

Figure 5.41: *The three gene repressilator with individual gene sites clustered together at the origin:* $g_1 = (0\mu m, 0.5\mu m, 0\mu m), g_2 = (0.5\cos(\pi/6)\mu m, -0.5\sin(\pi/6)\mu m, 0\mu m)$ *and* $g_3 = (-0.5\cos(\pi/6)\mu m, -0.5\sin(\pi/6)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

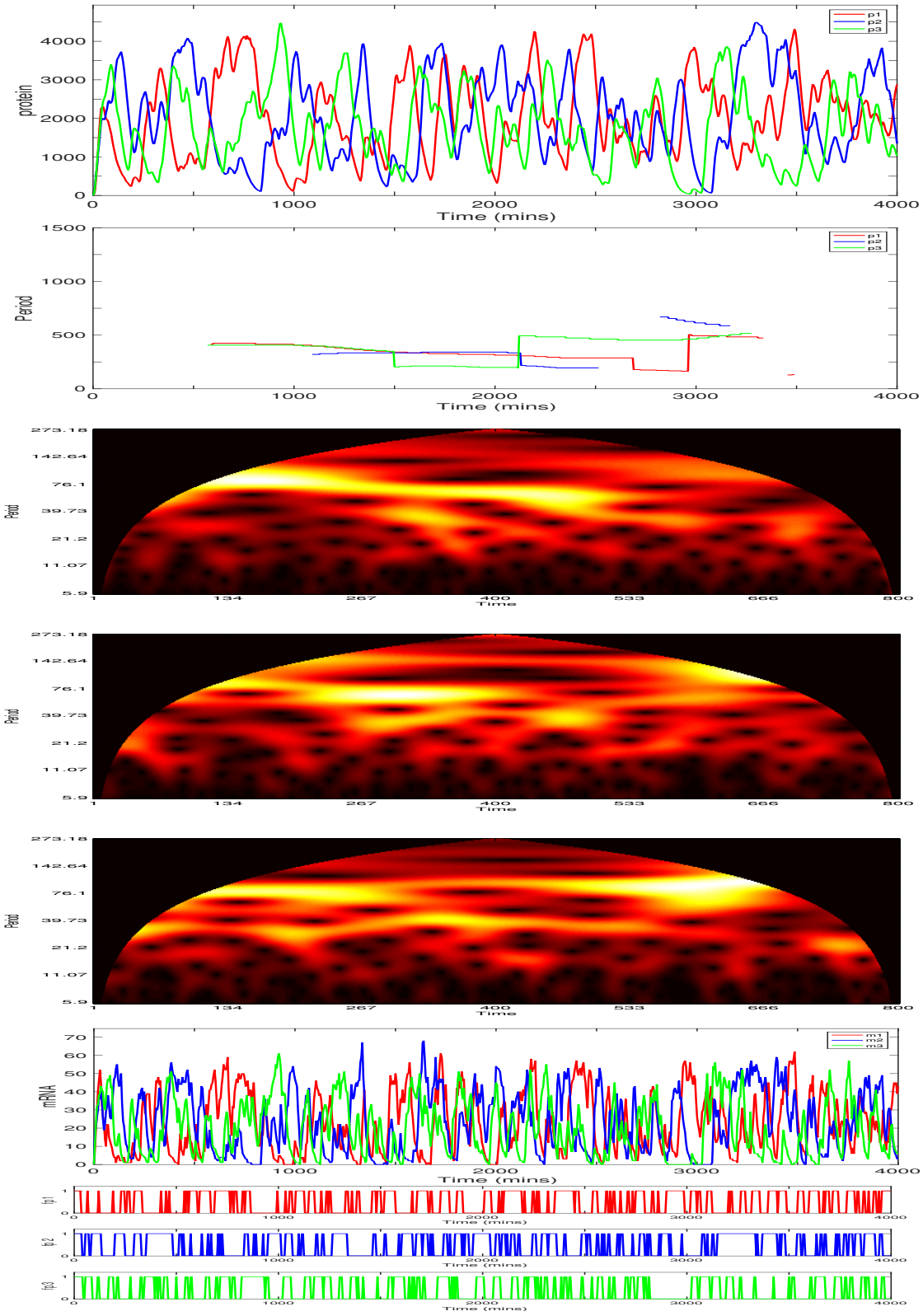

Figure 5.42: *The three gene repressilator with individual gene sites clustered together at the origin:* $g_1 = (0\mu m, 0.5\mu m, 0\mu m), g_2 = (0.5\cos(\pi/6)\mu m, -0.5\sin(\pi/6)\mu m, 0\mu m)$ *and* $g_3 = (-0.5\cos(\pi/6)\mu m, -0.5\sin(\pi/6)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

For $D = 1 \times 10^{-12} m^2 min^{-1}$ the average period for 100 trajectories was $\bar{T} = 595$, ranging between, $227 \leq \bar{T} \leq 1756$. The average protein number was $\bar{p} = 158$, ranging between $106 \leq \bar{p} \leq 200$. The average peak value of protein number was $\bar{p}_{peak} = 450$, ranging between $333 \leq \bar{p}_{peak} \leq 624$.

For $D = 1 \times 10^{-10} m^2 min^{-1}$ the average period for 100 trajectories was $\bar{T} = 394$, ranging between, $230 \leq \bar{T} \leq 848$. The average protein number was $\bar{p} = 1944$, ranging between $1592 \leq \bar{p} \leq 2327$. The average peak value of protein number was $\bar{p}_{peak} = 4649$, ranging between $3636 \leq \bar{p}_{peak} \leq 5776$.

Once again, the increase of the diffusion parameter, D by a magnitude of two, produces an increase in the average and peak protein numbers by a magnitude of one. In the smaller diffusion regime because there are less protein synthesised, there are less protein in the vicinity of its target gene. However, it does just take one single protein to bind to a target promoter, to induce a negative feedback. The probability increases for protein-promoter binding the more protein there is available in the voxel containing the target gene. This probability increases with the diffusion parameter as the protein levels increase also. For both diffusion regimes, each individual species can be compared to that of the one-gene repressilator. Oscillatory behaviour of the lower diffusion regime is a bit sketchy. With average periods ranging between 3.75 and 30 hours. There are a lot of stochastic fluctuations in degradation that add to the mix. Promoter sites seem to be turned off a lot of them time. Thus adding to the greater periods seen throughout the lower diffusion regime in comparison to the higher diffusion regime. For the higher diffusion regime, oscillatory behaviour is a lot more defined. The faster diffusion regime has average periods that range between 3.75 and 14 hours. We suggest that the increase of the diffusion parameter in a stochastic model can drive oscillations. The oscillations recorded are out of phase with one another and the amplitudes of each protein can differ also. The periods detected can sometimes run in conjunction with each other, however, because of the high stochastic nature,

the periods of each species can somewhat vary from each other. These finding are in somewhat agreement to the model of Elowitz & Leibler (2000). Their stochastic model showed species oscillating out of phase with each other and with differing amplitudes. We also suggest that maybe future work could investigate the different modes of the system and how this affects dynamics.

We now consider the gene site locations to be $g_1 = (x_1, y_1, z_1) = (0\mu m, 2.85\mu m, 0\mu m)$, $g_2 = (x_2, y_2, z_2) = (2.85\cos(\pi/6)\mu m, -2.85\sin(\pi/6)\mu m, 0\mu m)$ and $g_3 = (x_3, y_3, z_3) = (-2.85\cos(\pi/6)\mu m, -2.85\sin(\pi/6)\mu m, 0\mu m)$ for gene 1, 2 and 3 respectively i.e equidistant apart along the positive and negative x-axis and close to the nuclear membrane.

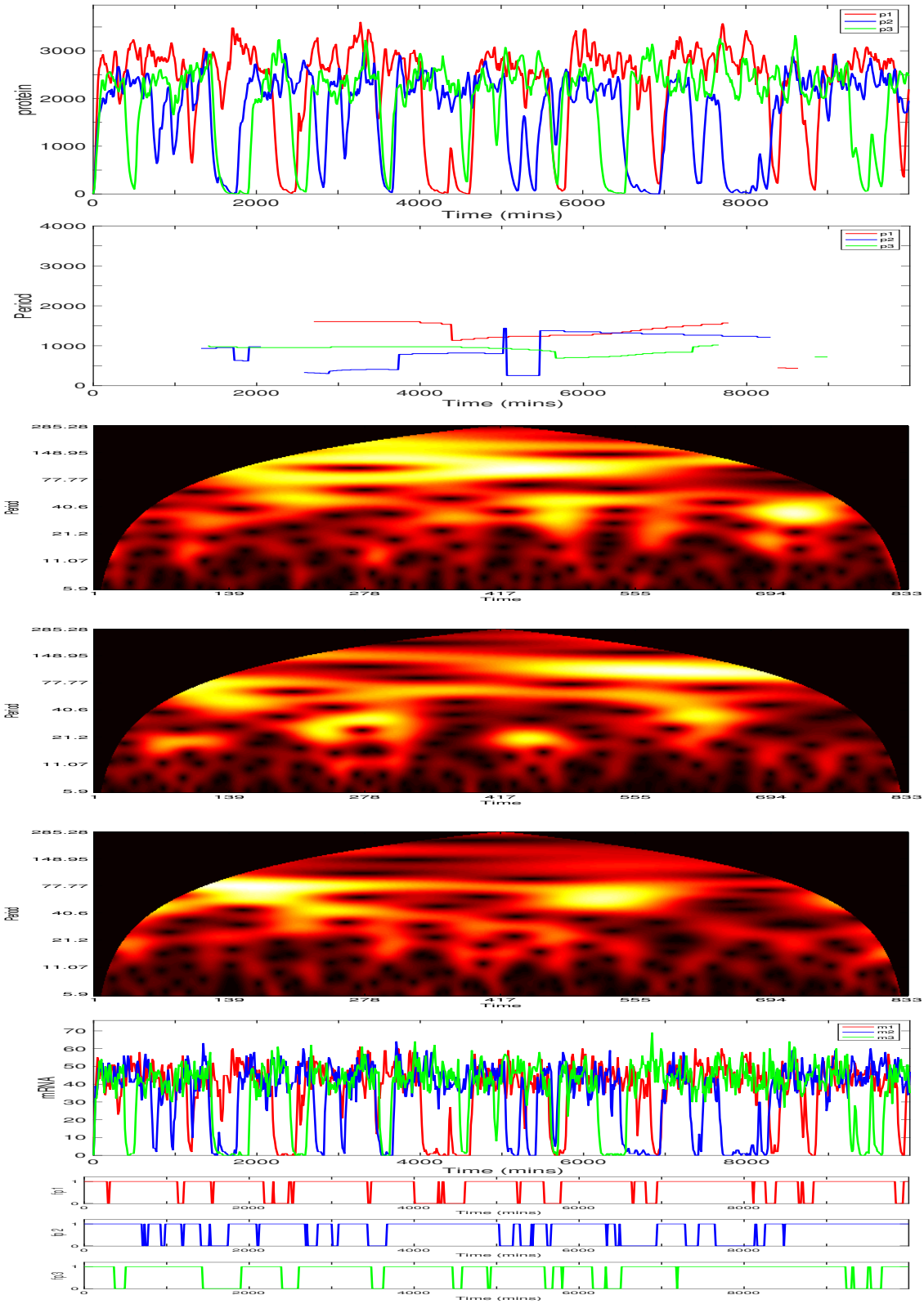

Figure 5.43: *The three gene repressilator with individual gene sites equidistant apart near the nuclear membrane:* $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (2.85\cos(\pi/6)\mu m, -2.85\sin(\pi/6)\mu m, 0\mu m)$ *and* $g_3 = (-2.85\cos(\pi/6)\mu m, -2.85\sin(\pi/6)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 12 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 12. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

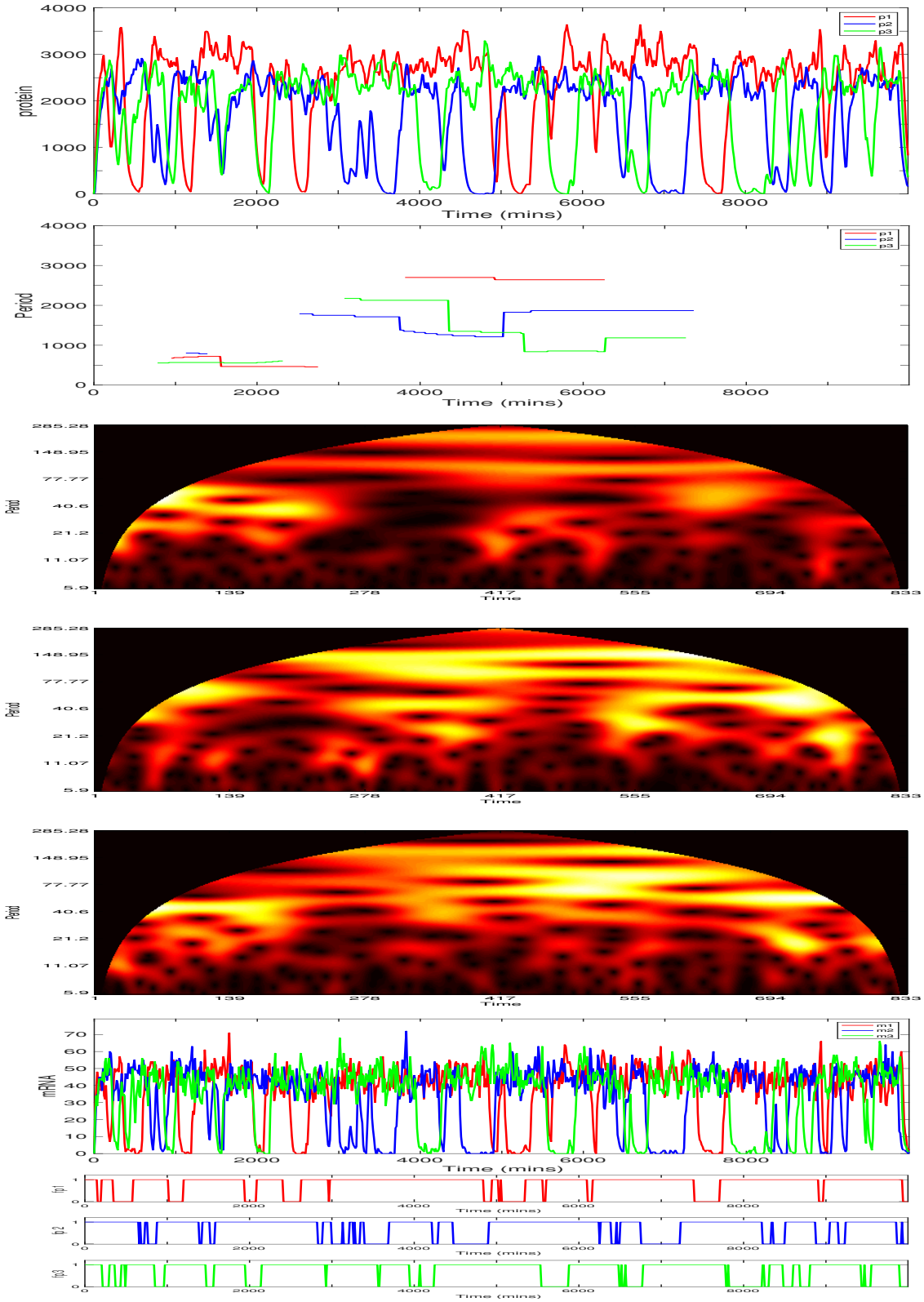

Figure 5.44: *The three gene repressilator with individual gene sites equidistant apart near the nuclear membrane:* $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (2.85\cos(\pi/6)\mu m, -2.85\sin(\pi/6)\mu m, 0\mu m)$ *and* $g_3 = (-2.85\cos(\pi/6)\mu m, -2.85\sin(\pi/6)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 12 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 12. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

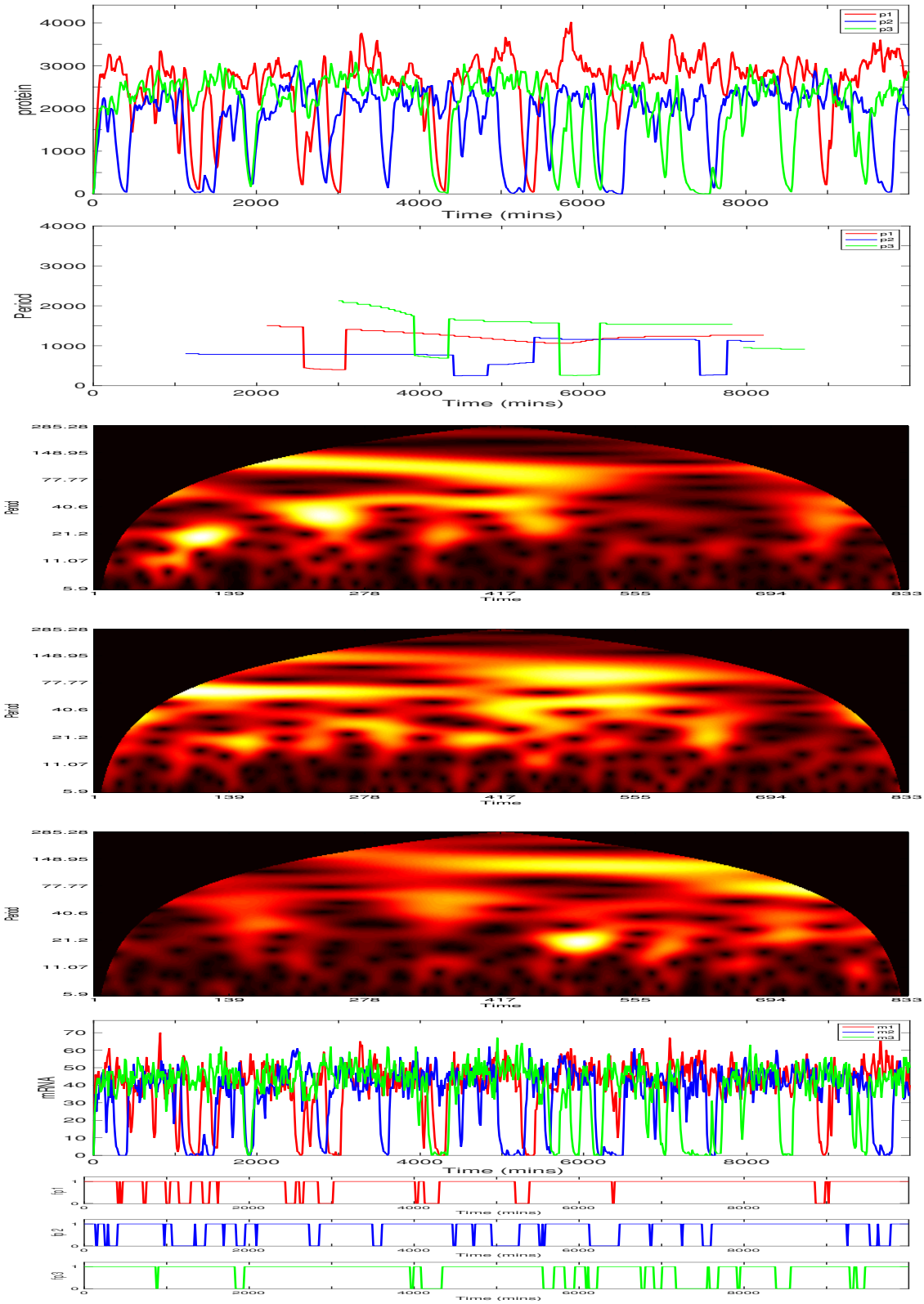

Figure 5.45: *The three gene repressilator with individual gene sites equidistant apart near the nuclear membrane:* $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (2.85\cos(\pi/6)\mu m, -2.85\sin(\pi/6)\mu m, 0\mu m)$ *and* $g_3 = (-2.85\cos(\pi/6)\mu m, -2.85\sin(\pi/6)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 12 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 12. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

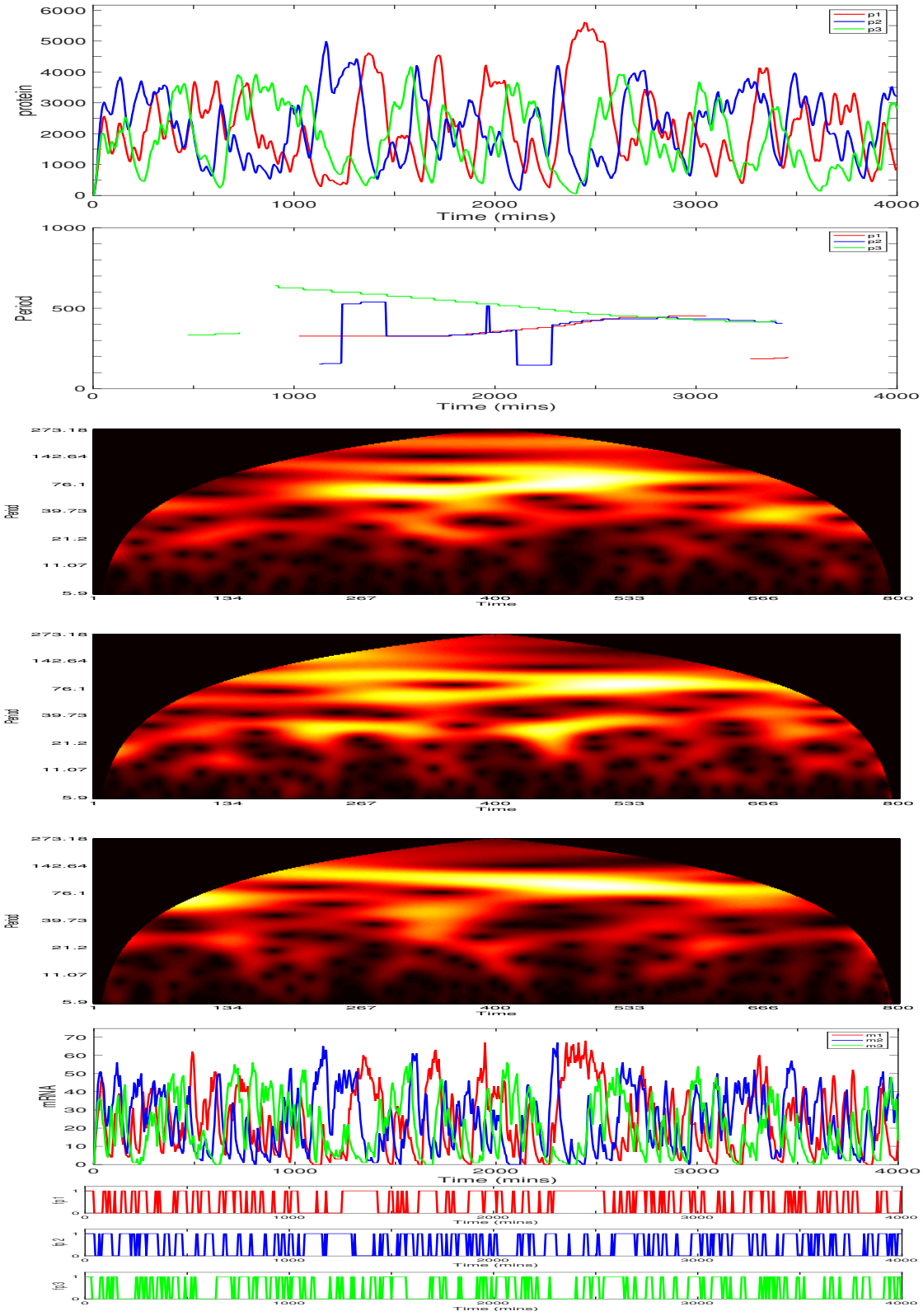

Figure 5.46: *The three gene repressilator with individual gene sites equidistant apart near the nuclear membrane:* $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (2.85\cos(\pi/6)\mu m, -2.85\sin(\pi/6)\mu m, 0\mu m)$ *and* $g_3 = (-2.85\cos(\pi/6)\mu m, -2.85\sin(\pi/6)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

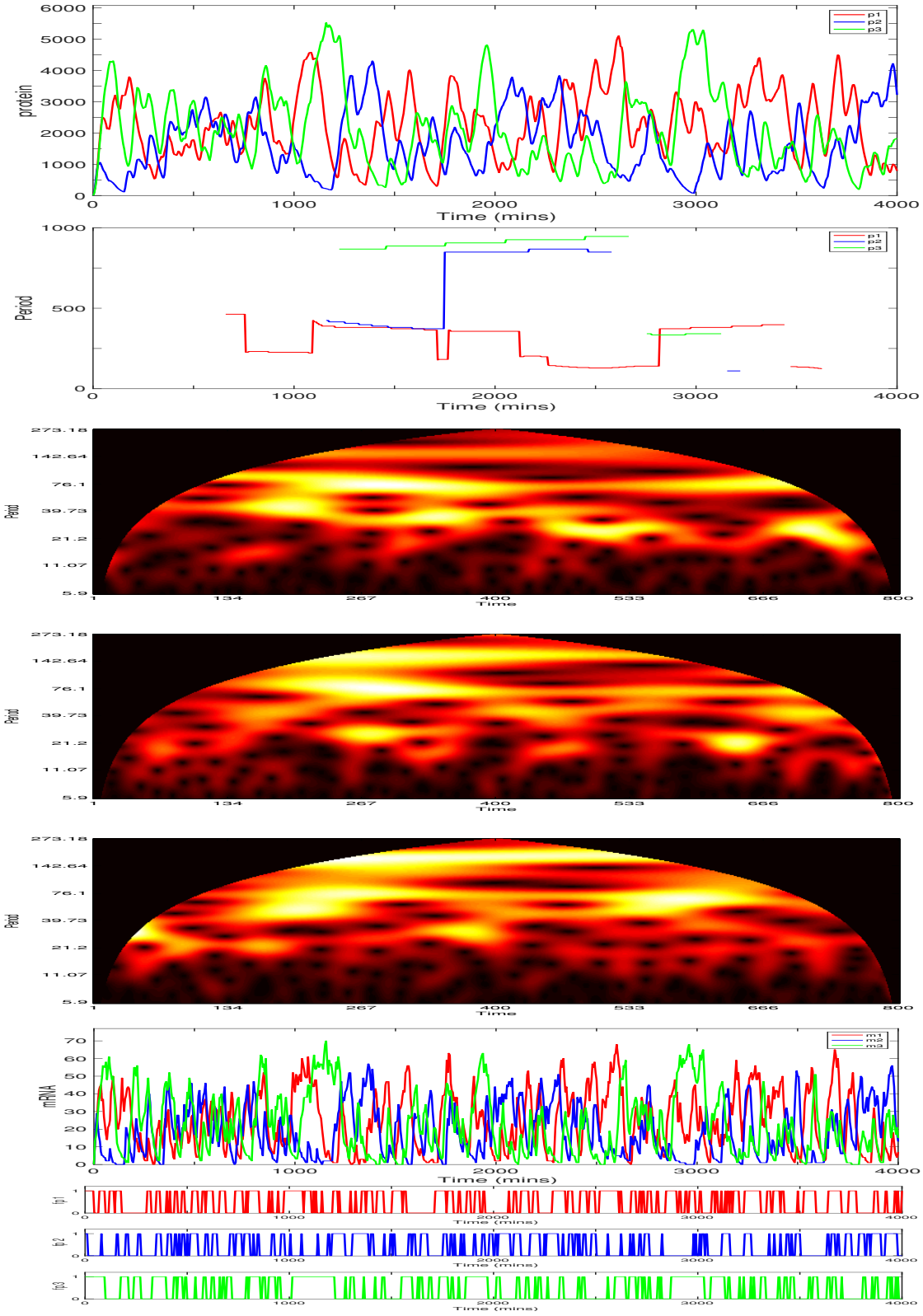

Figure 5.47: *The three gene repressilator with individual gene sites equidistant apart near the nuclear membrane:* $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (2.85\cos(\pi/6)\mu m, -2.85\sin(\pi/6)\mu m, 0\mu m)$ *and* $g_3 = (-2.85\cos(\pi/6)\mu m, -2.85\sin(\pi/6)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

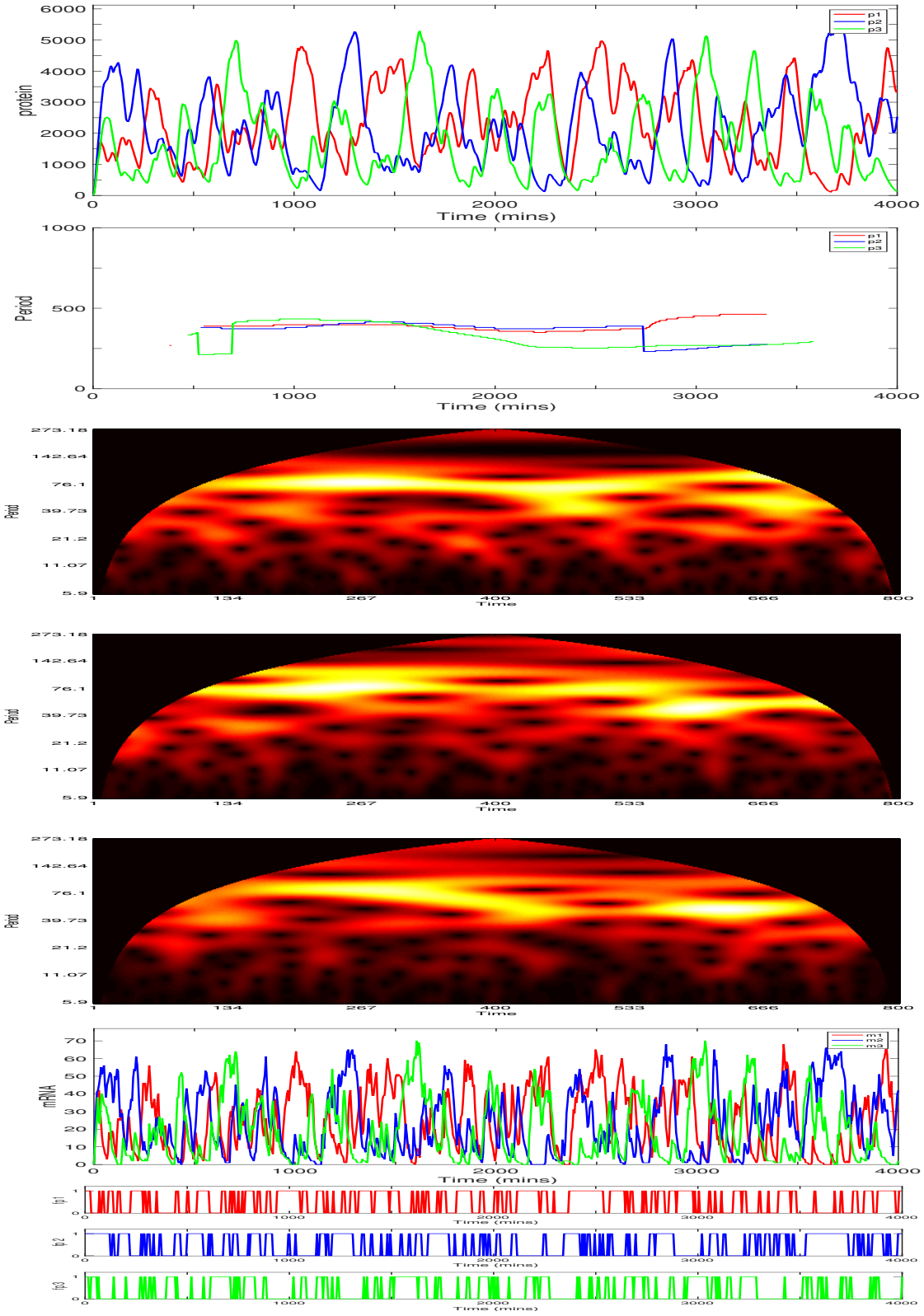

Figure 5.48: *The three gene repressilator with individual gene sites equidistant apart near the nuclear membrane:* $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (2.85\cos(\pi/6)\mu m, -2.85\sin(\pi/6)\mu m, 0\mu m)$ *and* $g_3 = (-2.85\cos(\pi/6)\mu m, -2.85\sin(\pi/6)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

For $D = 1 \times 10^{-12} m^2 min^{-1}$ the average period for 100 trajectories was $\bar{T} = 968$, ranging between, $521 \leq \bar{T} \leq 2533$. The average protein number was $\bar{p} = 2223$, ranging between $1709 \leq \bar{p} \leq 2624$. The average peak value of protein number was $\bar{p}_{peak} = 3735$, ranging between $3467 \leq \bar{p}_{peak} \leq 4355$.

For $D = 1 \times 10^{-10} m^2 min^{-1}$ the average period for 100 trajectories was $\bar{T} = 400$, ranging between, $223 \leq \bar{T} \leq 1077$. The average protein number was $\bar{p} = 2017$, ranging between $1645 \leq \bar{p} \leq 2449$. The average peak value of protein number was $\bar{p}_{peak} = 4830$, ranging between $3888 \leq \bar{p}_{peak} \leq 5695$.

For gene sites at opposite ends of the x-axis, situated close to the nuclear membrane, there is a similarity to the two-gene repressilator. Each gene site is maximally away from the other two gene sites and equidistant away from the origin. For the smaller diffusion regime, each protein can maximally be expressed due to the fact that most of the other protein species are located on a small radius of proximity around their own promoters and thus, a smaller likelihood of negative feedback will occur due to the infrequent promoter binding. However, since the distance to each of their target genes is less than that of the case for the two-gene repressilator, there is a higher probability that a protein will travel toward and reach and bind to its target promoter. This can explain the average period calculated being 160 minutes less than the average 'period' calculated for the two-gene repressilator. We can claim here that for this case of gene site position in the lower diffusion regime, the three-gene repressilator is a stronger oscillator than the two-gene repressilator. For the faster diffusion regime, we can once again see diffusion aid in the onset of oscillations, as we see more defined oscillations.

We now consider the gene site locations to be $(x_1, y_1, z_1) = (0\mu m, 2.85\mu m, 0\mu m)$, $(x_2, y_2, z_2) = (0.5\mu m, 2.85 - \frac{\sqrt{3}}{2}\mu m, 0\mu m)$ and $(x_3, y_3, z_3) = (-0.5\mu m, 2.85 - \frac{\sqrt{3}}{2}\mu m, 0\mu m)$ for gene 1, 2 and 3 respectively i.e a gene sites are in a cluster close together and close to the nuclear membrane. This can be viewed as translating the origin cluster of gene

sites approximately $2\mu m$ along the positive y-axis.

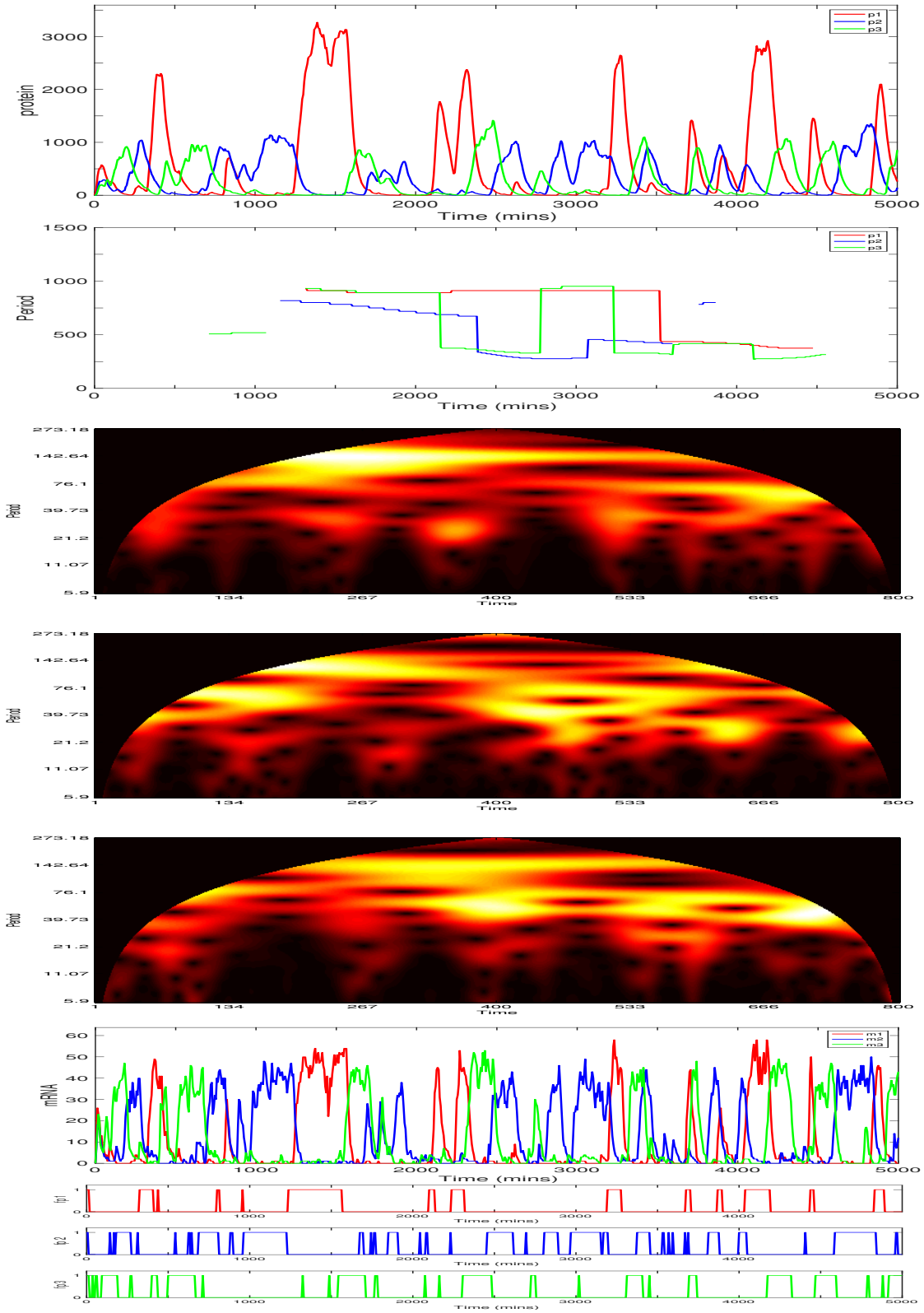

Figure 5.49: *The three gene repressilator with individual gene sites clustered together near the nuclear membrane:* $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (0.5\cos(\pi/6)\mu m, 2.1\mu m, 0\mu m)$ *and* $g_3 = (-0.5\cos(\pi/6)\mu m, 2.1\mu m, 0\mu m)$*. A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$*. Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 6.25 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 6.25. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

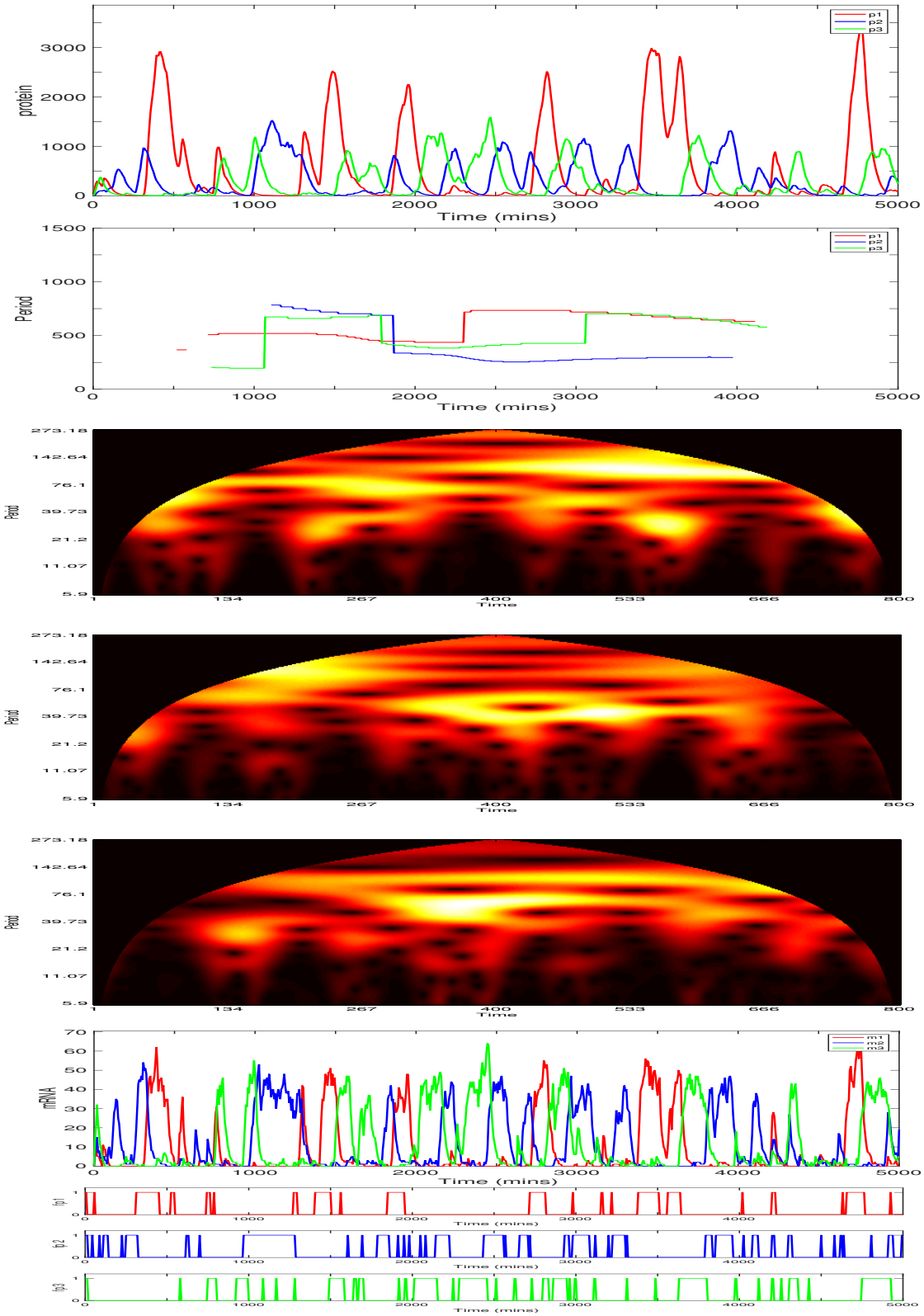

Figure 5.50: *The three gene repressilator with individual gene sites clustered together near the nuclear membrane: $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (0.5\cos(\pi/6)\mu m, 2.1\mu m, 0\mu m)$ and $g_3 = (-0.5\cos(\pi/6)\mu m, 2.1\mu m, 0\mu m)$. A diffusion coefficient $D = 1 \times 10^{-12} m^2 min^{-1}$. Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 6.25 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 6.25. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

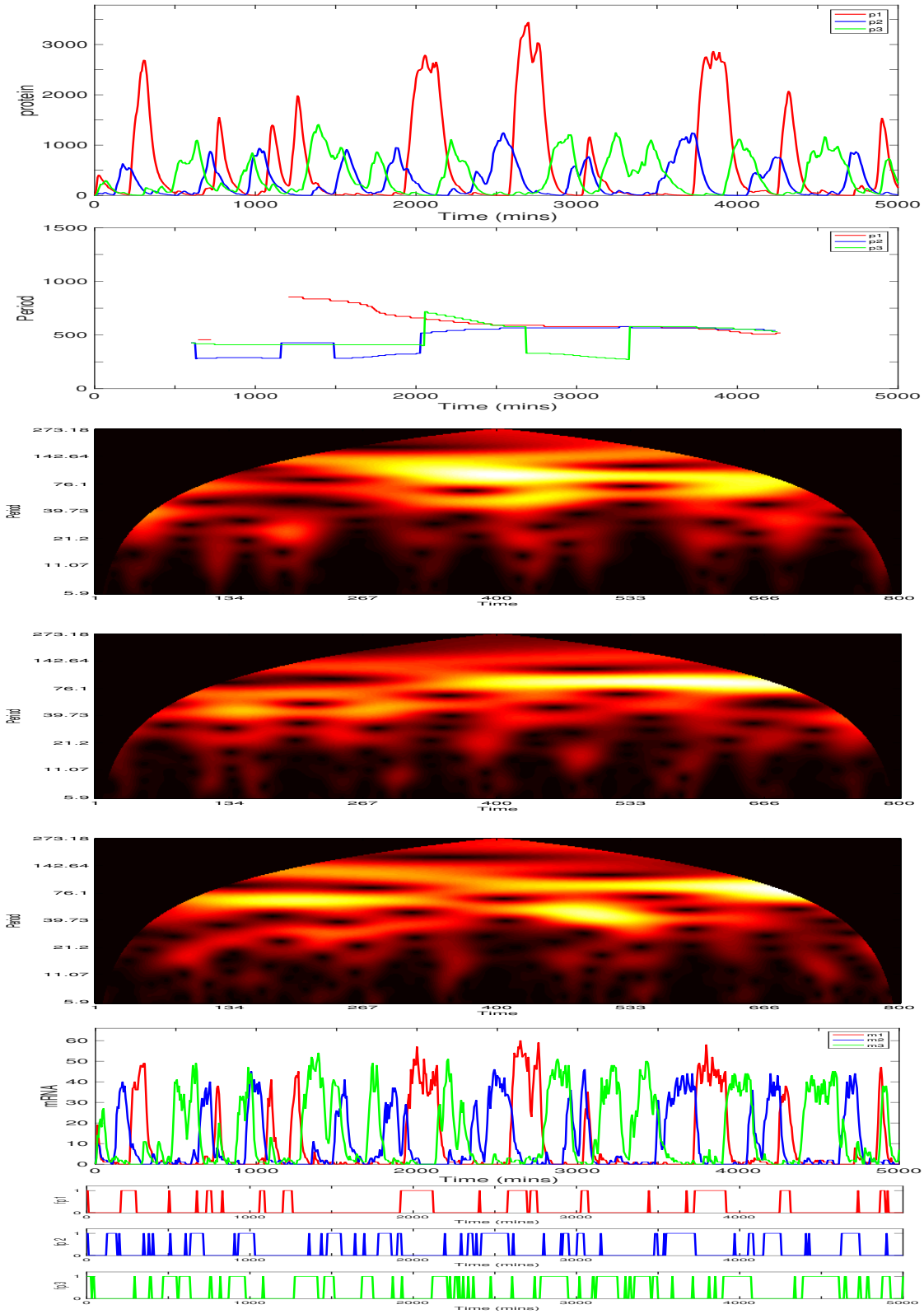

Figure 5.51: *The three gene repressilator with individual gene sites clustered together near the nuclear membrane: $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (0.5\cos(\pi/6)\mu m, 2.1\mu m, 0\mu m)$ and $g_3 = (-0.5\cos(\pi/6)\mu m, 2.1\mu m, 0\mu m)$. A diffusion coefficient $D = 1 \times 10^{-12} m^2 min^{-1}$. Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 6.25 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 6.25. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

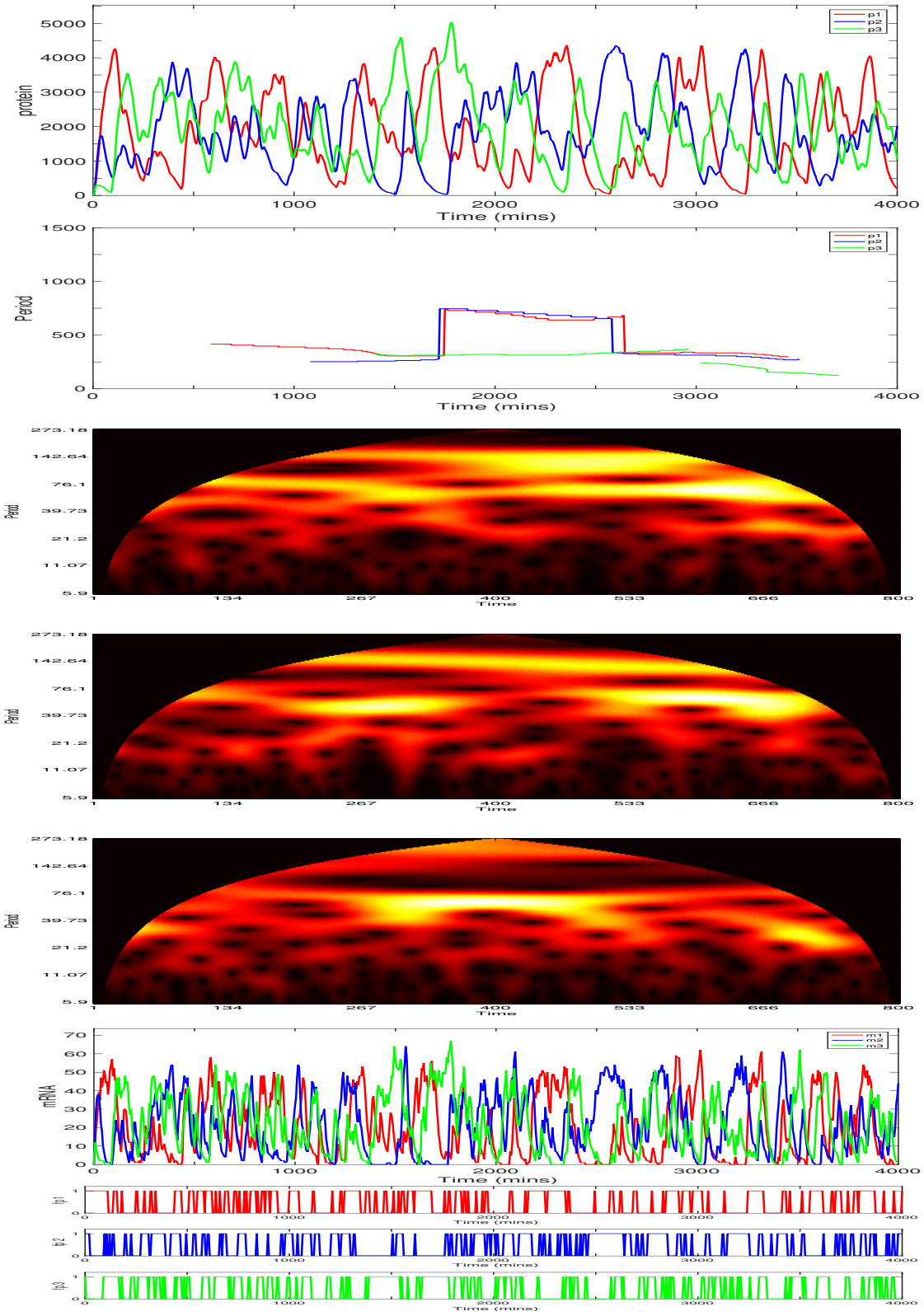

Figure 5.52: *The three gene repressilator with individual gene sites clustered together near the nuclear membrane:* $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (0.5\cos(\pi/6)\mu m, 2.1\mu m, 0\mu m)$ *and* $g_3 = (-0.5\cos(\pi/6)\mu m, 2.1\mu m, 0\mu m)$*. A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$*. Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

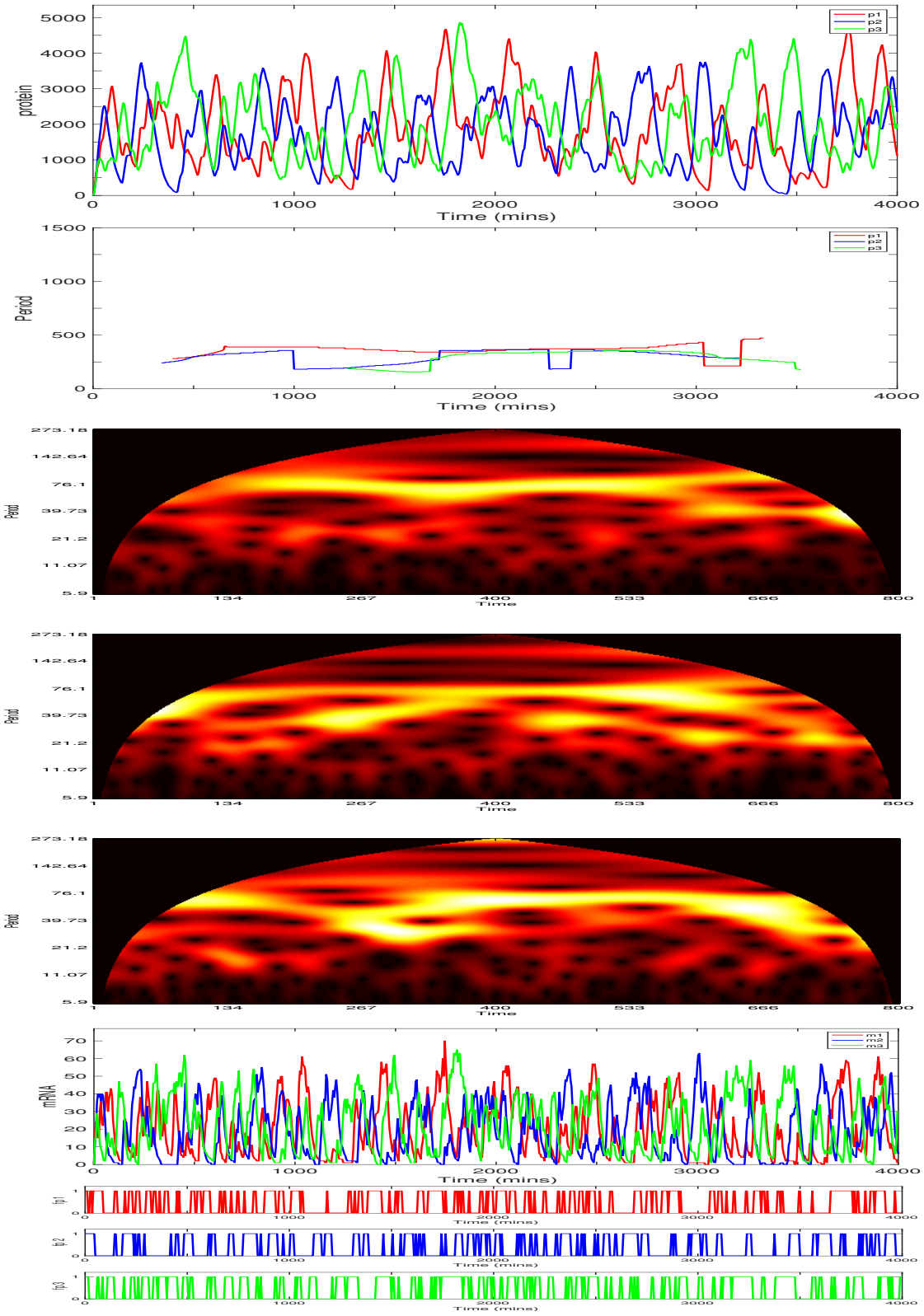

Figure 5.53: *The three gene repressilator with individual gene sites clustered together near the nuclear membrane:* $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (0.5\cos(\pi/6)\mu m, 2.1\mu m, 0\mu m)$ *and* $g_3 = (-0.5\cos(\pi/6)\mu m, 2.1\mu m, 0\mu m)$*. A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$*. Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

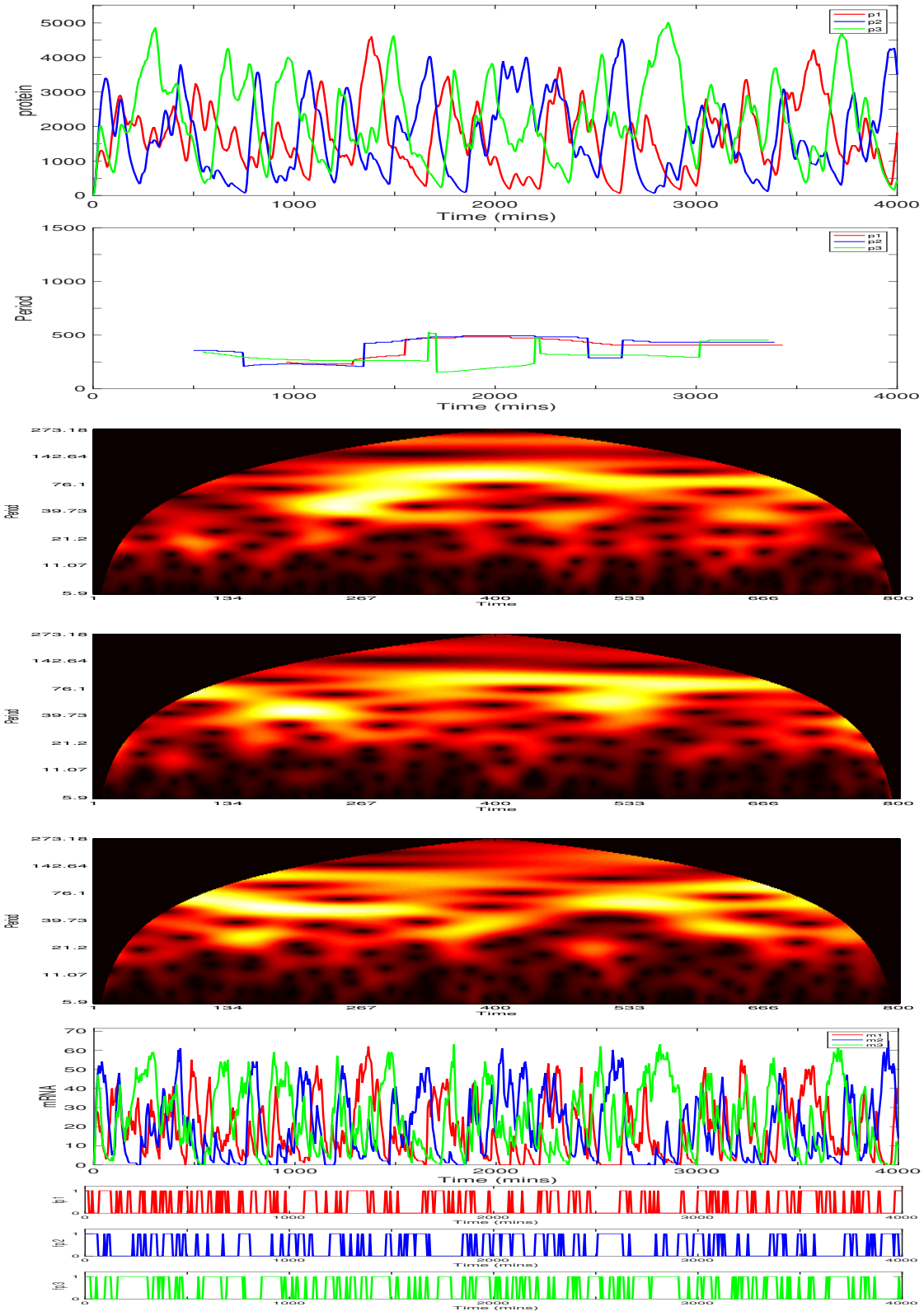

Figure 5.54: *The three gene repressilator with individual gene sites clustered together near the nuclear membrane:* $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (0.5\cos(\pi/6)\mu m, 2.1\mu m, 0\mu m)$ *and* $g_3 = (-0.5\cos(\pi/6)\mu m, 2.1\mu m, 0\mu m)$*. A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$*. Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products. and Gene3/products. Top to bottom: Protein, period, heat map of protein period 1, 2 and 3, mRNA and Free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

For $D = 1 \times 10^{-12} m^2 min^{-1}$ the average period for 200 trajectories was $\bar{T} = 540$, ranging between, $324 \leq \bar{T} \leq 954$. The average protein number for protein one was $\bar{p} = 449$, ranging between $208 \leq \bar{p} \leq 734$. The average peak value of protein number was $\bar{p}_{peak} = 3060$, ranging between $2058 \leq \bar{p}_{peak} \leq 3835$. The average protein number for protein two and three was $\bar{p} = 292$, ranging between $148 \leq \bar{p} \leq 420$. The average peak value level of proteins two and three was $\bar{p}_{peak} = 1339$, ranging between $1025 \leq \bar{p}_{peak} \leq 1763$.

For $D = 1 \times 10^{-10} m^2 min^{-1}$ the average period for 200 trajectories was $\bar{T} = 400$, ranging between, $241 \leq \bar{T} \leq 790$. The average protein number was $\bar{p} = 2024$, ranging between $1710 \leq \bar{p} \leq 2495$. The average peak value of protein number was $\bar{p}_{peak} = 4771$, ranging between $4215 \leq \bar{p}_{peak} \leq 5610$.

For the increase in D the average peak protein number of protein one increased by approximately 150%. However the average protein number of the other two increased by 400%. For proteins two and three, as D is increased to the faster diffusion regime, we see an increase in their levels of a magnitude of one.

The larger amount of protein one can account for its position within the gene cluster. Gene site one is closer to the nuclear membrane than the other two gene sites by approximately $0.6\mu m$. This positioning is more optimal for more mRNA to locate and translate protein one in the cytoplasm. In this scheme, oscillations are out of phase and peak amplitudes vary over time also. For the faster diffusion regime, the difference in gene site positions does not have much affect. Again, the increase in diffusion aided in the onset of oscillations.

See figure 5.55 for a summary of the three-gene repressilator. For the faster diffusion regime, the three-gene repressilator behaves similarly to the one-gene repressilator. However, the periods of oscillation detected were about twice as long on average.

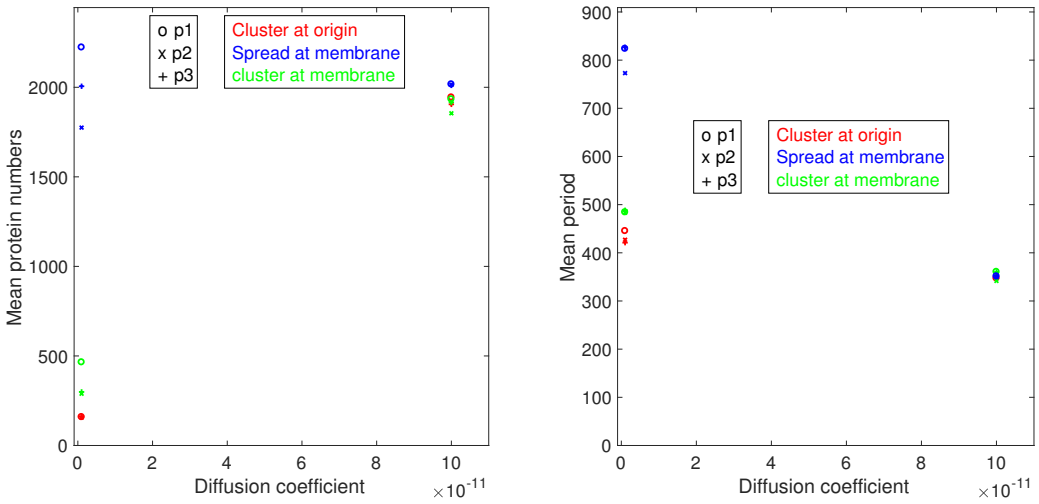

Figure 5.55: *Left: The mean value of protein calculated for the three-gene repressilator against the diffusion parameter. Right: The mean period calculated for the three-gene repressilator agains the diffusion parameter. Red denotes the gene site cluster at the origin, blue denotes the gene sites at the membrane on opposite ends of the x-axis and green denotes the cluster of genes close to the nuclear membrane.*

5.4 The four gene repressilator

We now consider the spatial-stochastic model for the four-gene repressilator. Thus we add another gene into our ring GRN, such that gene one indirectly represses gene two; gene two indirectly represses gene three; gene three indirectly represses gene four and gene four indirectly represses gene one.

We first consider the gene site locations to be positions: $g_1 = (x_1, y_1, z_1) = (0\mu m, \frac{1}{\sqrt{2}}\mu m, 0\mu m)$, $g_2 = (x_2, y_2, z_2) = (\frac{1}{\sqrt{2}}\mu m, 0\mu m, 0\mu m)$, $g_3 = (x_3, y_3, z_3) = (0\mu m, -\frac{1}{\sqrt{2}}\mu m, 0\mu m)$ and $g_4 = (x_4, y_4, z_4) = (-\frac{1}{\sqrt{2}}\mu m, 0\mu m, 0\mu m)$ for gene 1, 2, 3 and 4 respectively i.e each gene lies on the vertex of a square with sides of length $1\mu m$. We term this position, a cluster of gene sites at the origin.

The first set of three simulations are for the smaller diffusion regime, $D = 1 \times 10^{-12}$ and the second set of three simulations are for the faster diffusion regime, $D = 1 \times 10^{-10}$.

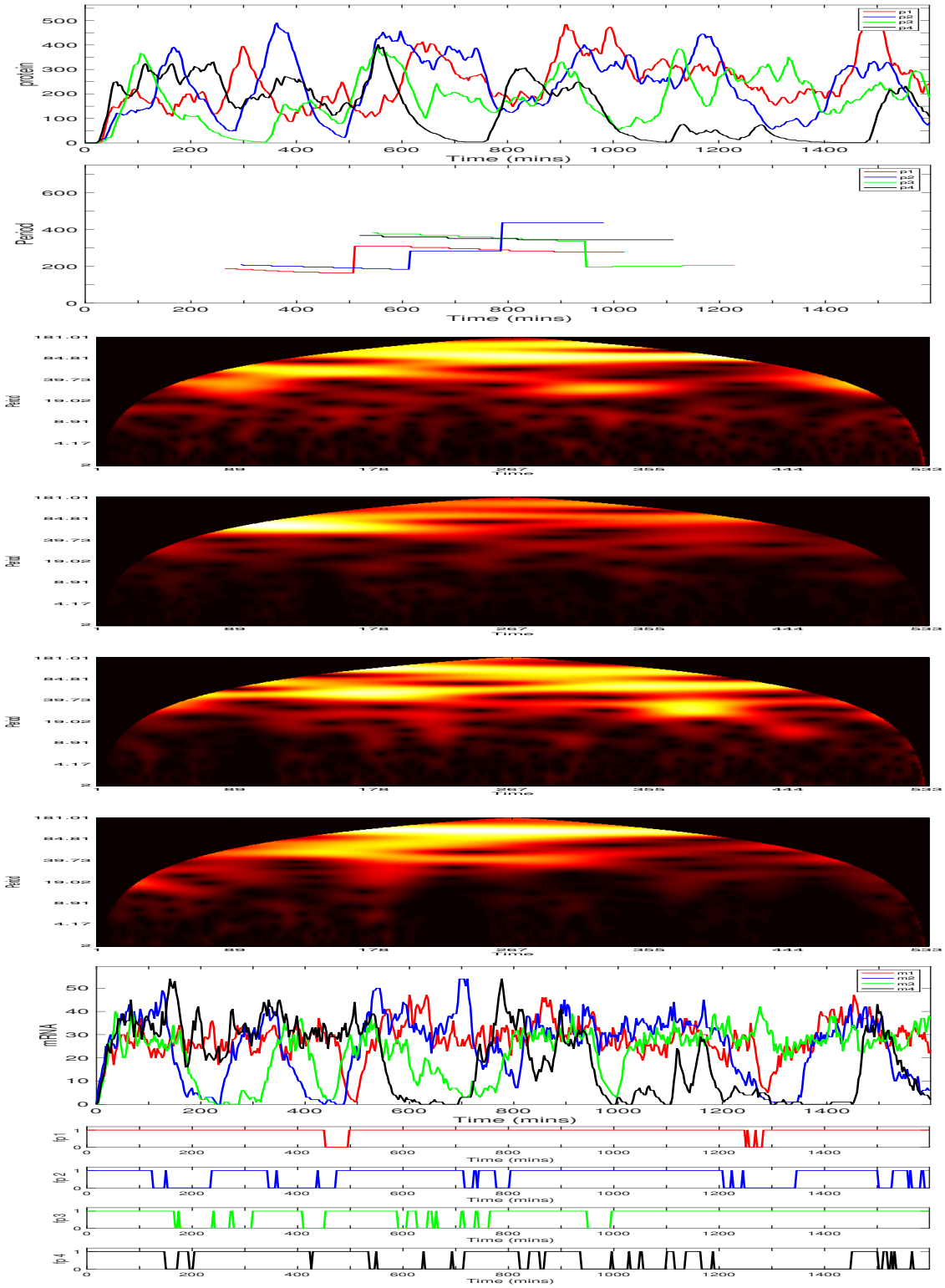

Figure 5.56: *The four gene repressilator with individual gene sites clustered together at the origin:* $g_1 = (0\mu m, \frac{1}{\sqrt{2}}\mu m, 0\mu m), g_2 = (\frac{1}{\sqrt{2}}\mu m, 0\mu m, 0\mu m)$, $g_3 = (0\mu m, -\frac{1}{\sqrt{2}}\mu m, 0\mu m)$ *and* $g_4 = (-\frac{1}{\sqrt{2}}\mu m, 0\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products. and Gene4/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

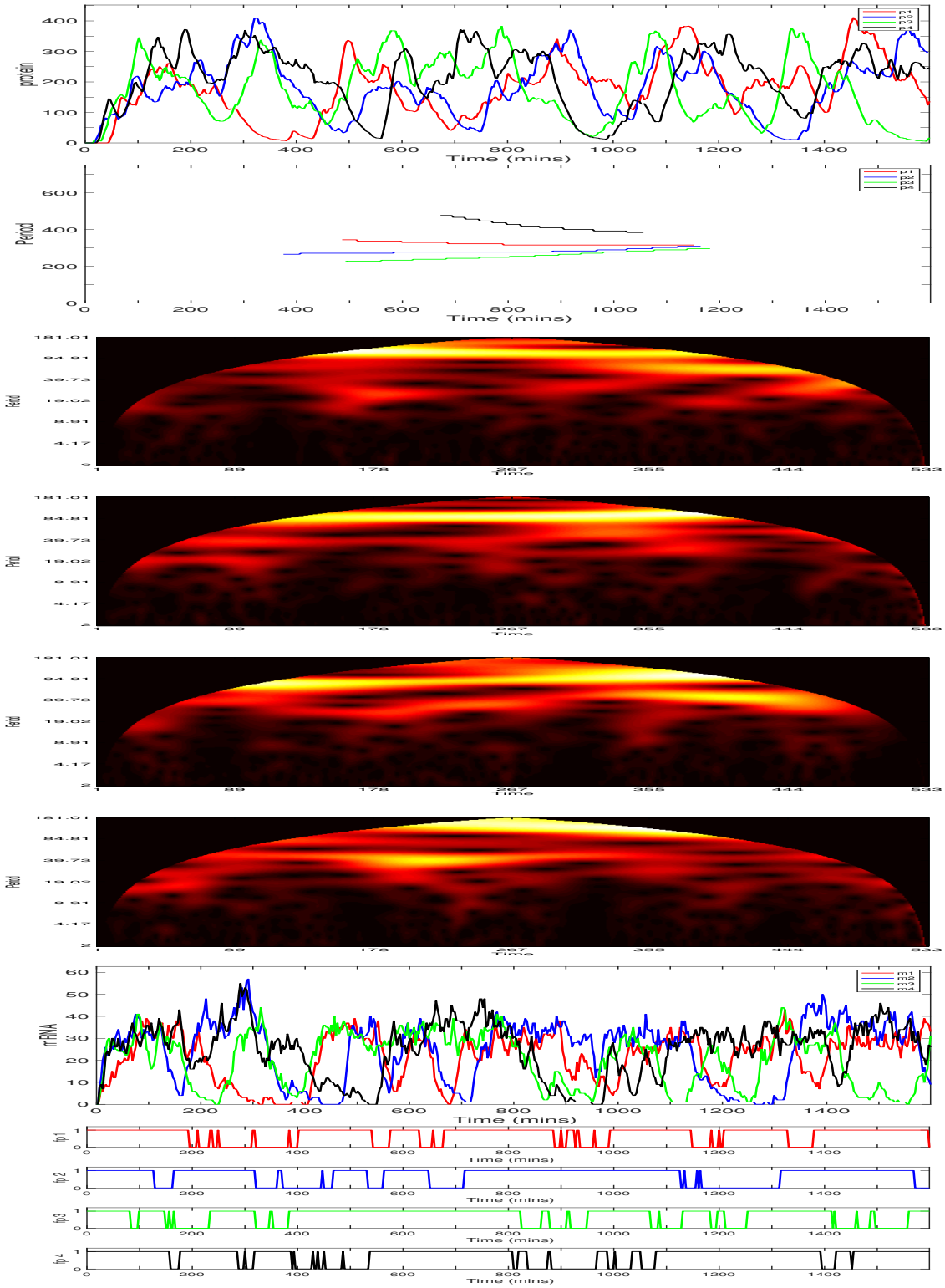

Figure 5.57: *The four gene repressilator with individual gene sites clustered together at the origin:* $g_1 = (0\mu m, \frac{1}{\sqrt{2}}\mu m, 0\mu m), g_2 = (\frac{1}{\sqrt{2}}\mu m, 0\mu m, 0\mu m)$, $g_3 = (0\mu m, -\frac{1}{\sqrt{2}}\mu m, 0\mu m)$ *and* $g_4 = (-\frac{1}{\sqrt{2}}\mu m, 0\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products. and Gene4/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

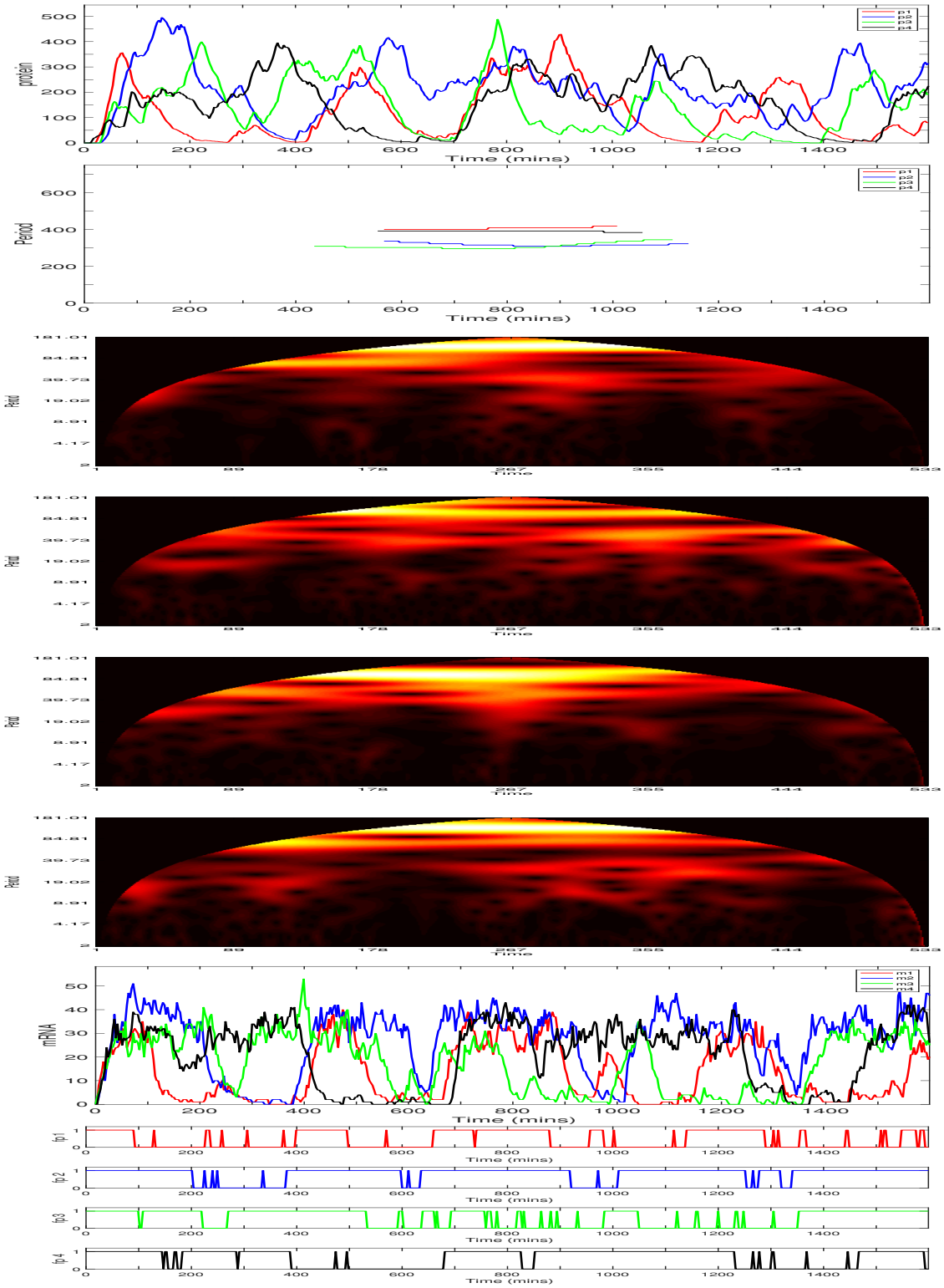

Figure 5.58: *The four gene repressilator with individual gene sites clustered together at the origin:* $g_1 = (0\mu m, \frac{1}{\sqrt{2}}\mu m, 0\mu m), g_2 = (\frac{1}{\sqrt{2}}\mu m, 0\mu m, 0\mu m)$, $g_3 = (0\mu m, -\frac{1}{\sqrt{2}}\mu m, 0\mu m)$ *and* $g_4 = (-\frac{1}{\sqrt{2}}\mu m, 0\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products. and Gene4/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

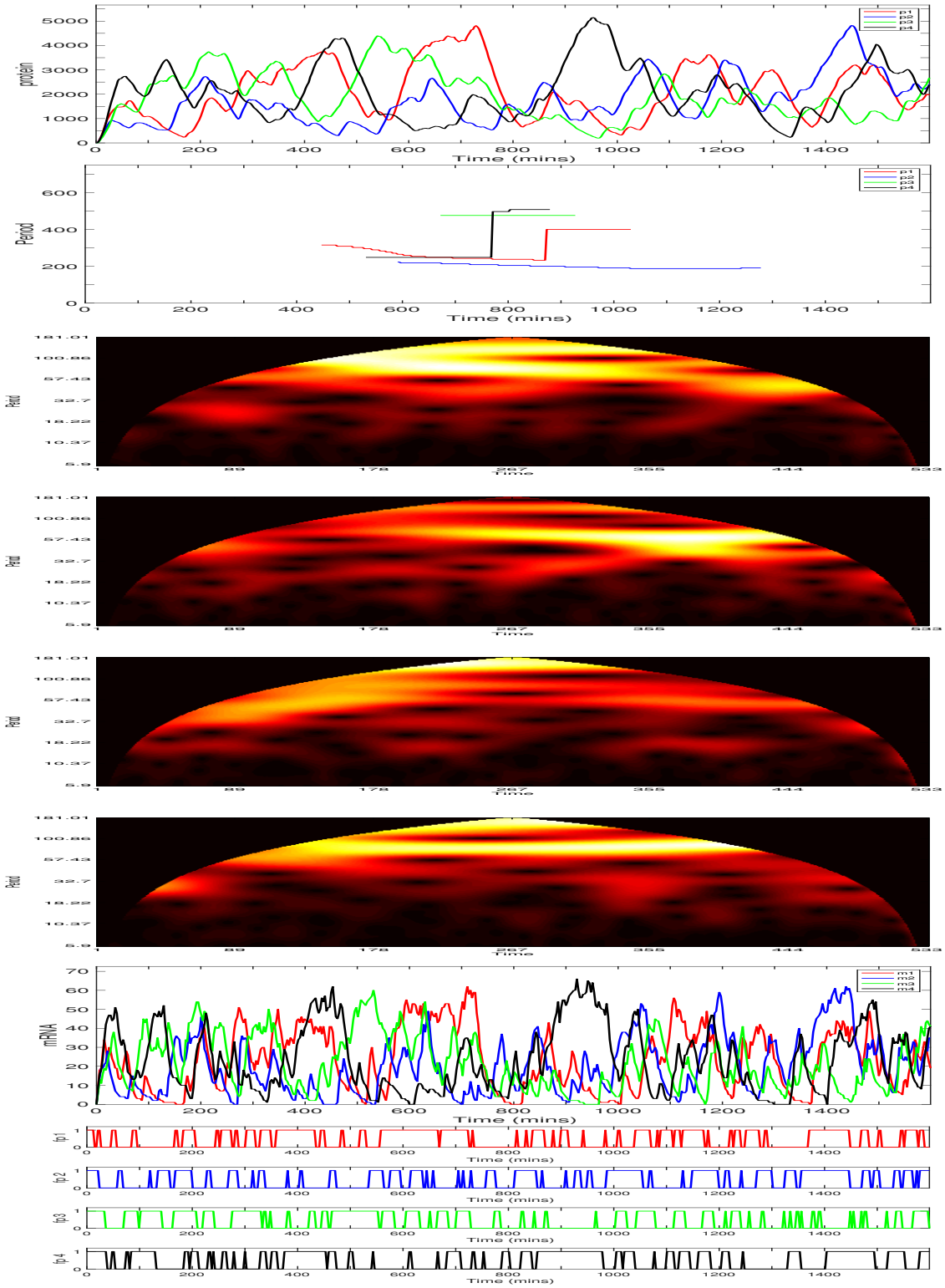

Figure 5.59: *The four gene repressilator with individual gene sites clustered together at the origin:* $g_1 = (0\mu m, \frac{1}{\sqrt{2}}\mu m, 0\mu m), g_2 = (\frac{1}{\sqrt{2}}\mu m, 0\mu m, 0\mu m)$, $g_3 = (0\mu m, -\frac{1}{\sqrt{2}}\mu m, 0\mu m)$ *and* $g_4 = (-\frac{1}{\sqrt{2}}\mu m, 0\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products. and Gene4/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

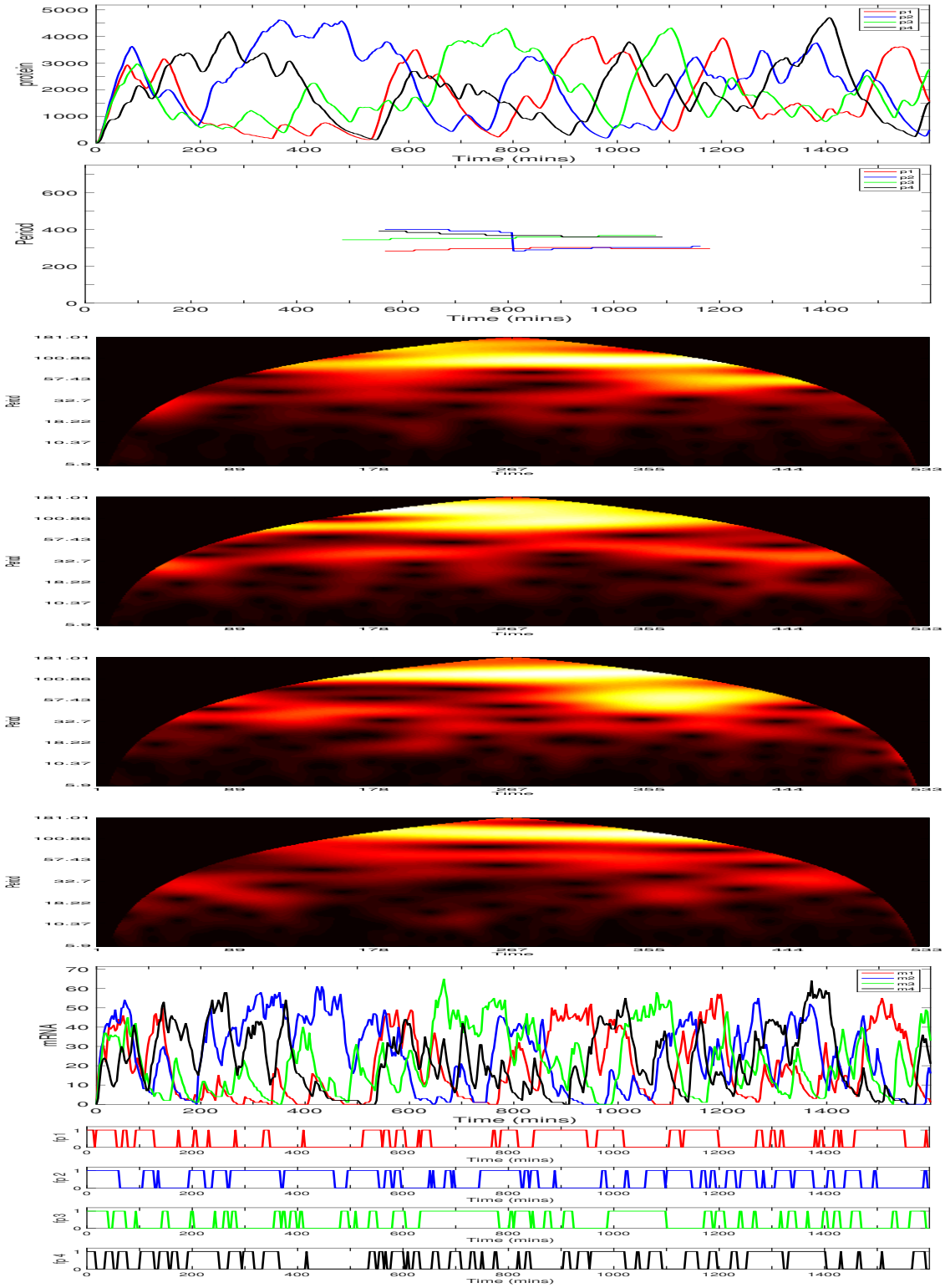

Figure 5.60: *The four gene repressilator with individual gene sites clustered together at the origin:* $g_1 = (0\mu m, \frac{1}{\sqrt{2}}\mu m, 0\mu m), g_2 = (\frac{1}{\sqrt{2}}\mu m, 0\mu m, 0\mu m)$, $g_3 = (0\mu m, -\frac{1}{\sqrt{2}}\mu m, 0\mu m)$ *and* $g_4 = (-\frac{1}{\sqrt{2}}\mu m, 0\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products. and Gene4/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

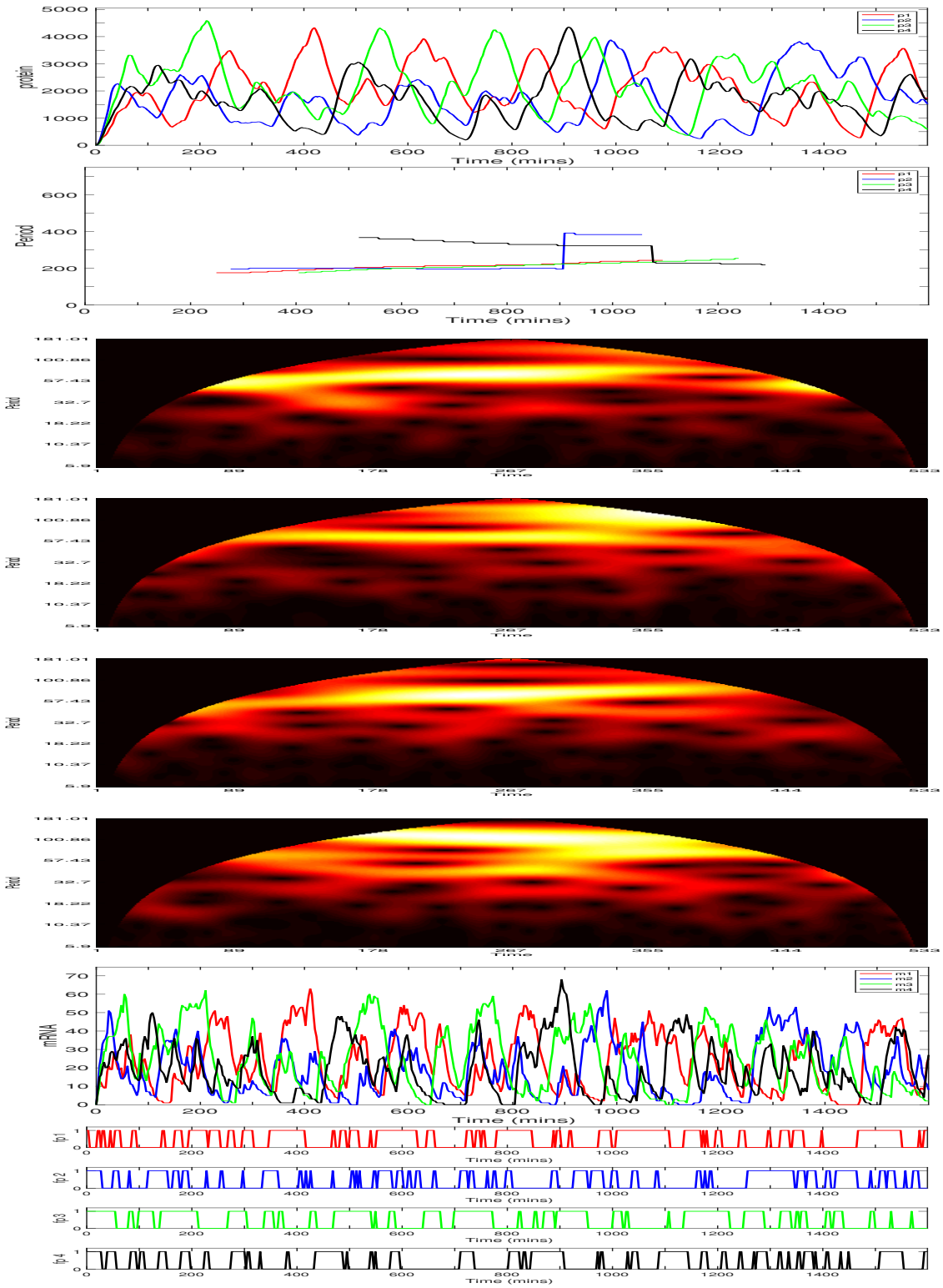

Figure 5.61: *The four gene repressilator with individual gene sites clustered together at the origin:* $g_1 = (0\mu m, \frac{1}{\sqrt{2}}\mu m, 0\mu m), g_2 = (\frac{1}{\sqrt{2}}\mu m, 0\mu m, 0\mu m)$, $g_3 = (0\mu m, -\frac{1}{\sqrt{2}}\mu m, 0\mu m)$ *and* $g_4 = (-\frac{1}{\sqrt{2}}\mu m, 0\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products. and Gene4/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

For $D = 1 \times 10^{-12} m^2 min^{-1}$ the average period for 100 trajectories was $T = 350$, ranging between, $127 \leq T \leq 555$. The average protein number was $p_{av} = 190$, ranging between $107 \leq p_{av} \leq 287$. The average peak value of protein number was $p_{peak} = 481$, ranging between $307 \leq p_{av} \leq 729$.

For $D = 1 \times 10^{-10} m^2 min^{-1}$ the average period for 100 trajectories was $T = 290$, ranging between, $112 \leq T \leq 555$. The average protein number was $p_{av} = 1913$, ranging between $1235 \leq p_{av} \leq 2792$. The average peak value of protein number was $p_{peak} = 4239$, ranging between $2995 \leq p_{av} \leq 5357$.

As before, the magnitude in protein levels increases by a magnitude of one as the diffusion is increased. This is due to the larger amount of mRNA able to reach the cytoplasm and translate its corresponding mRNA. Thus, for the faster diffusion regime, there is a higher level of $protein_{i-1}$-$promoter_i$ binding and unbinding. This can explain the lower periods calculated for the faster diffusion regime in comparison to the slower diffusion regime. However, oscillations are a bit elusive, particularly for the slower diffusion regime, although some of the heat maps show defined modal periods. We suspect that if we ran the simulations for a longer time, higher modal periods would appear. For the time frame we have used the shorter period modes have been detected, leading for us to speculate that there are high and low modes of oscillation in the system. Each mode may be important for gene expression on different time scales: acting as a clock on more than one level. This could be a case for further investigation.

We now consider the gene site locations to be positions: $g_1 = (x_1, y_1, z_1) = (0\mu m, 2.85\mu m, 0\mu m)$, $g_2 = (x_2, y_2, z_2) = (2.85\mu m, 0\mu m, 0\mu m)$, $g_3 = (x_3, y_3, z_3) = (0\mu m, -2.85\mu m, 0\mu m)$ and $g_4 = (x_4, y_4, z_4) = (-2.85\mu m, 0\mu m, 0\mu m)$ for gene 1, 2, 3 and 4 respectively i.e each gene lies on the vertex of a square with sides of length $1\mu m$. We term this position, a cluster of gene sites at the origin.

The first set of three simulations are for the smaller diffusion regime, $D = 1 \times 10^{-12}$

and the second set of three simulations are for the faster diffusion regime, $D = 1 \times 10^{-10}$.

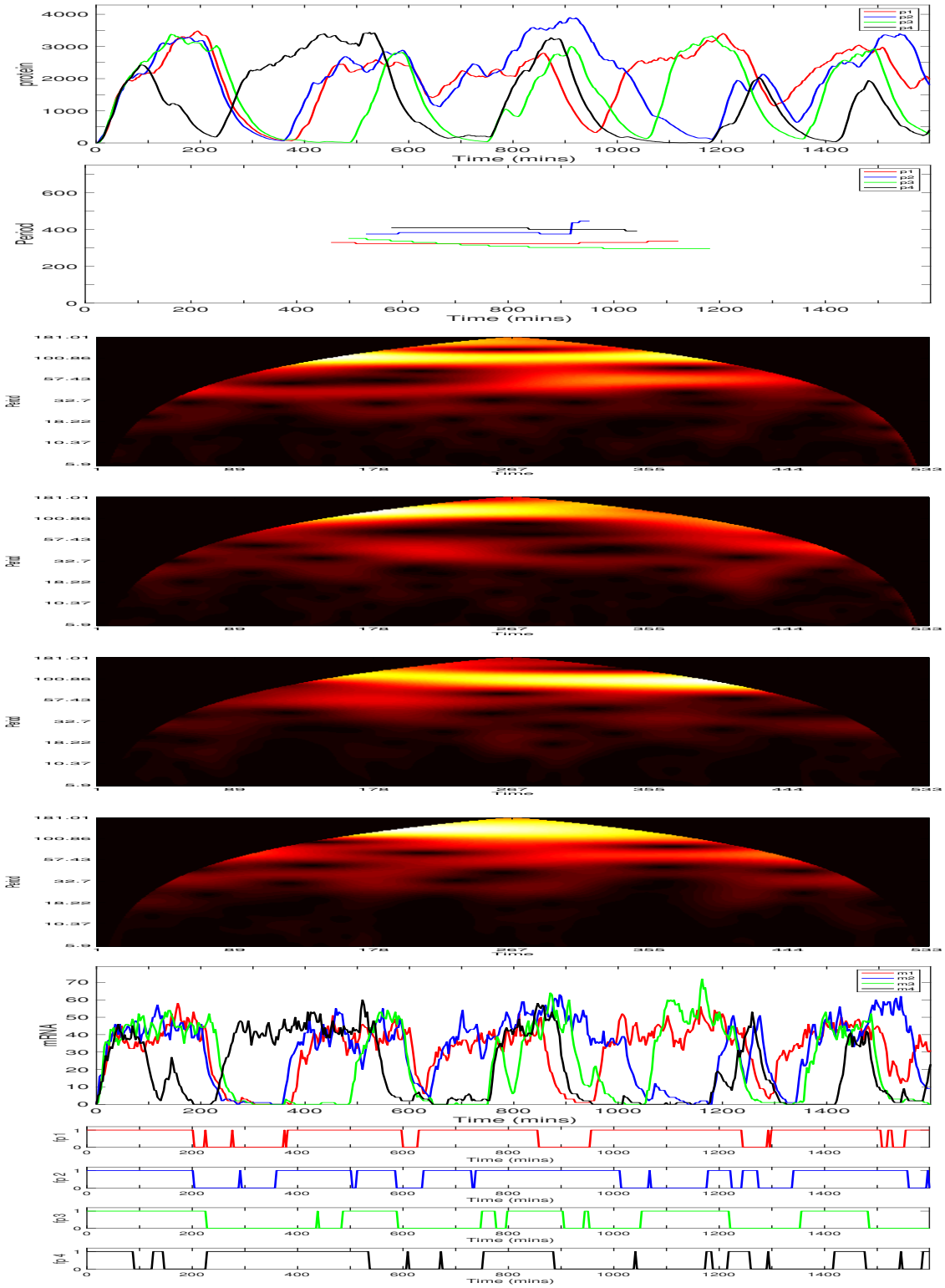

Figure 5.62: *The four gene repressilator with individual gene sites equidistant apart at the nuclear membrane:* $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (2.85\mu m, 0\mu m, 0\mu m)$, $g_3 = (0\mu m, -2.85\mu m, 0\mu m)$ *and* $g_4 = (-2.85\mu m, 0\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products. and Gene4/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

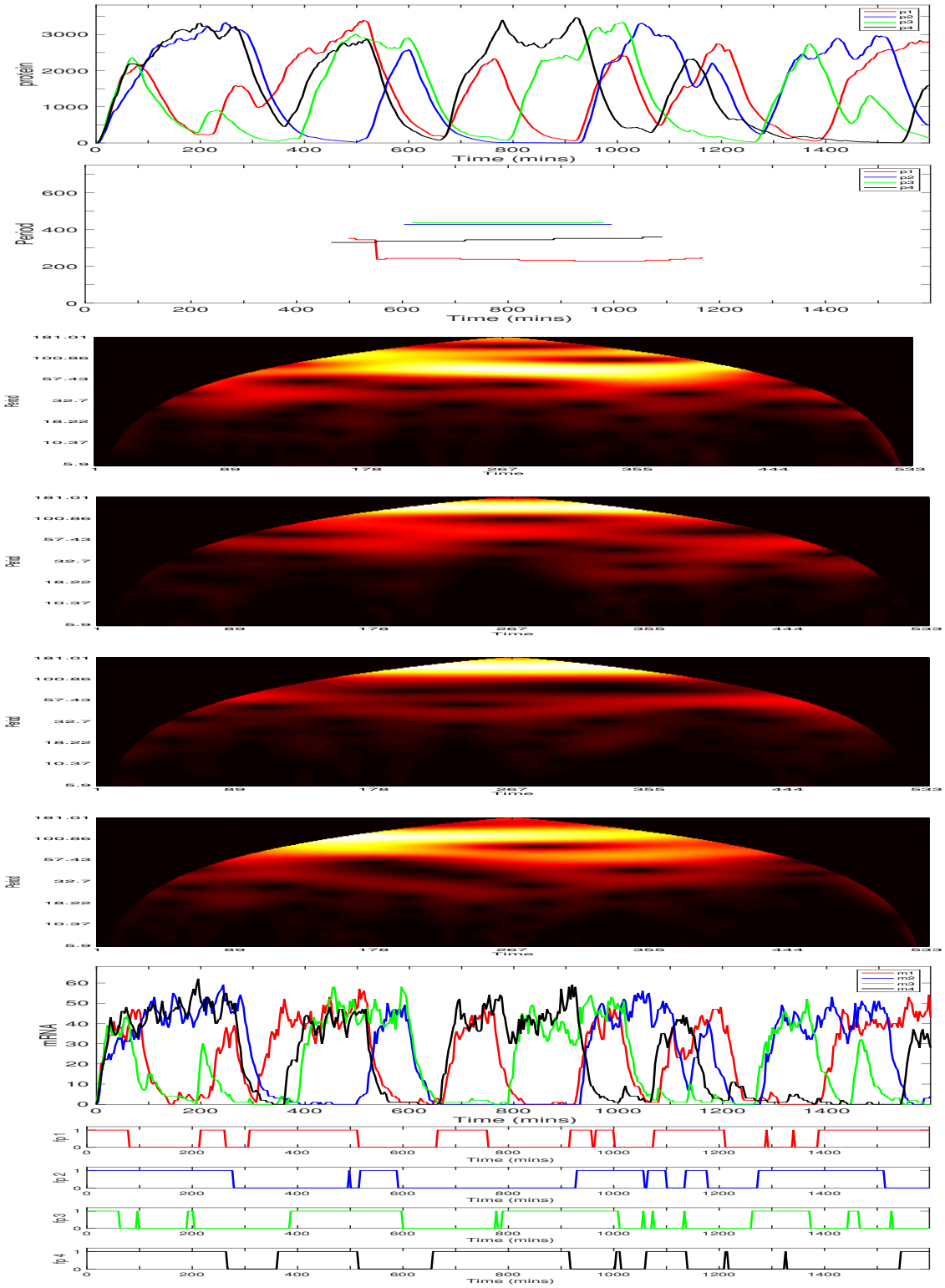

Figure 5.63: *The four gene repressilator with individual gene sites equidistant apart at the nuclear membrane:* $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (2.85\mu m, 0\mu m, 0\mu m)$, $g_3 = (0\mu m, -2.85\mu m, 0\mu m)$ *and* $g_4 = (-2.85\mu m, 0\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products. and Gene4/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

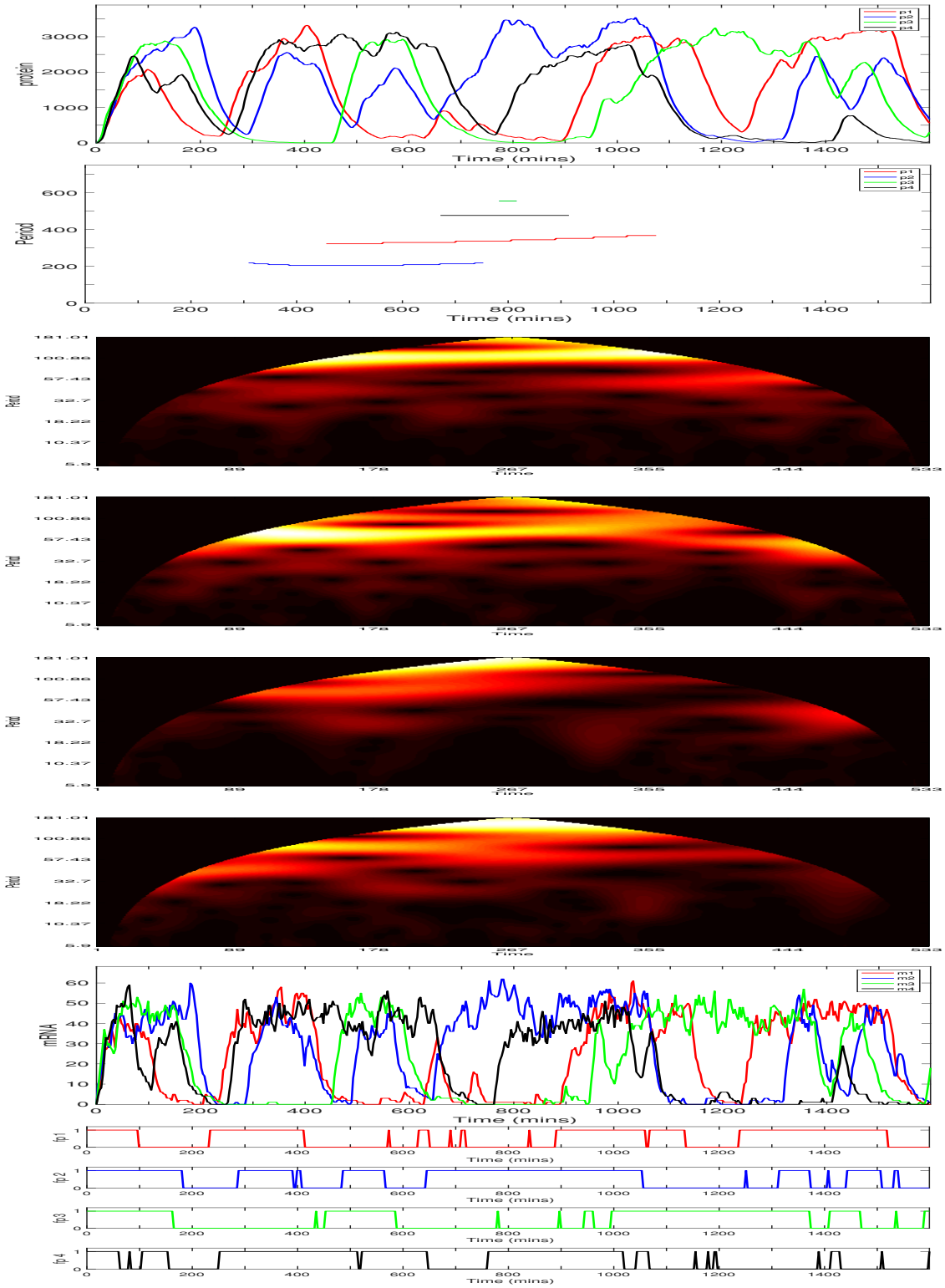

Figure 5.64: *The four gene repressilator with individual gene sites equidistant apart at the nuclear membrane:* $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (2.85\mu m, 0\mu m, 0\mu m)$, $g_3 = (0\mu m, -2.85\mu m, 0\mu m)$ *and* $g_4 = (-2.85\mu m, 0\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products. and Gene4/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

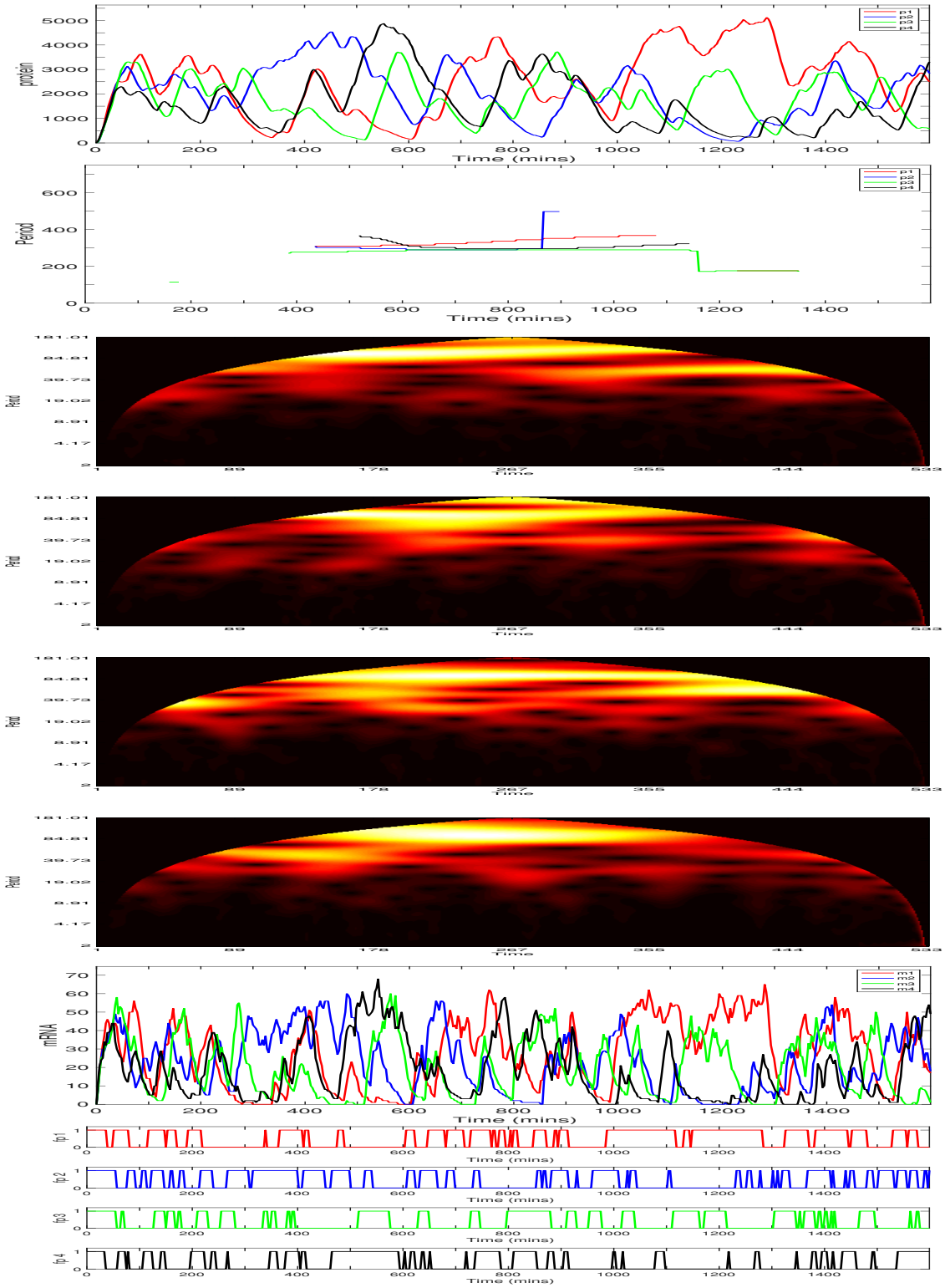

Figure 5.65: *The four gene repressilator with individual gene sites equidistant apart at the nuclear membrane:* $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (2.85\mu m, 0\mu m, 0\mu m)$, $g_3 = (0\mu m, -2.85\mu m, 0\mu m)$ *and* $g_4 = (-2.85\mu m, 0\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products. and Gene4/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

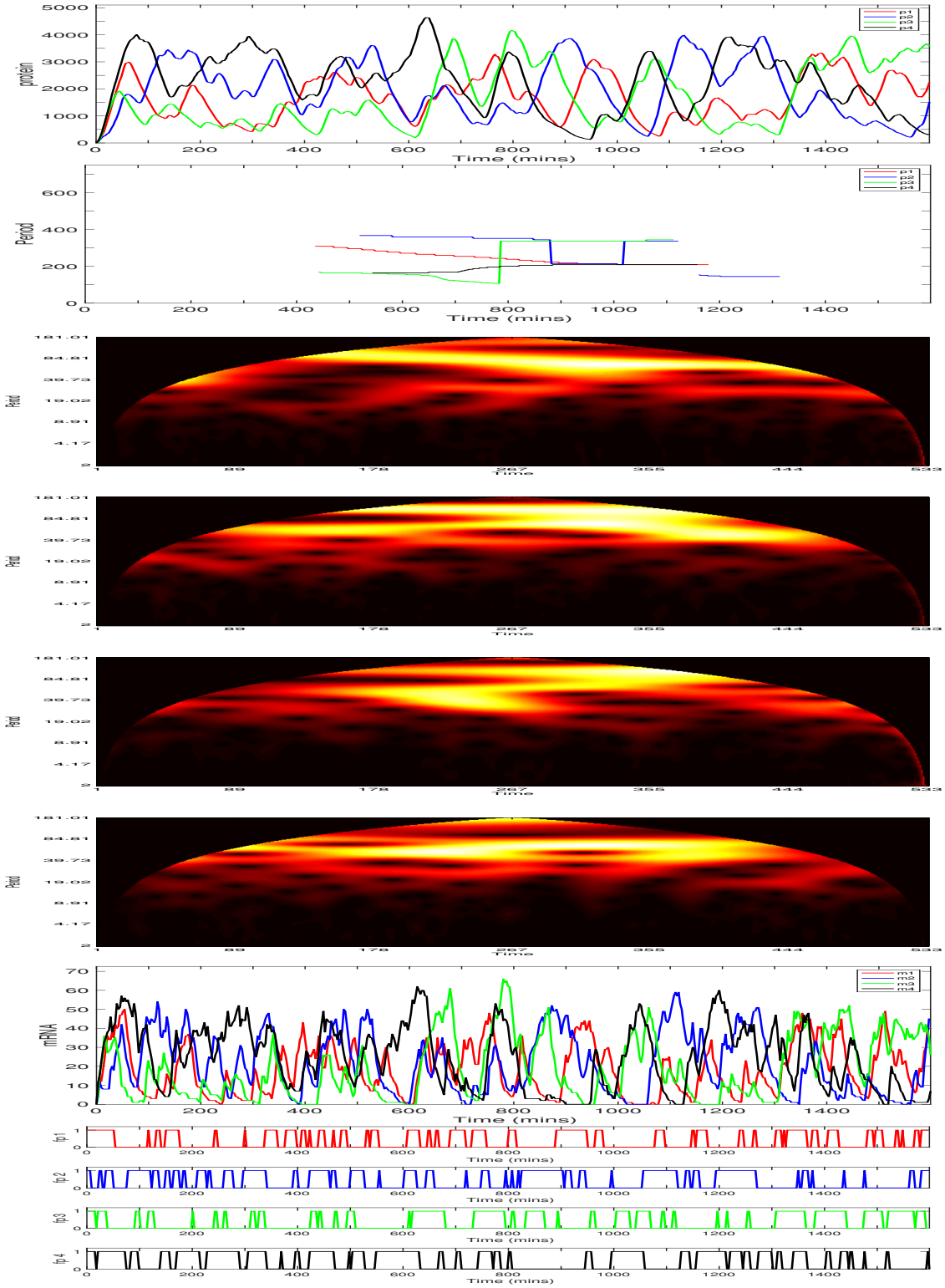

Figure 5.66: *The four gene repressilator with individual gene sites equidistant apart at the nuclear membrane:* $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (2.85\mu m, 0\mu m, 0\mu m)$, $g_3 = (0\mu m, -2.85\mu m, 0\mu m)$ *and* $g_4 = (-2.85\mu m, 0\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products. and Gene4/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

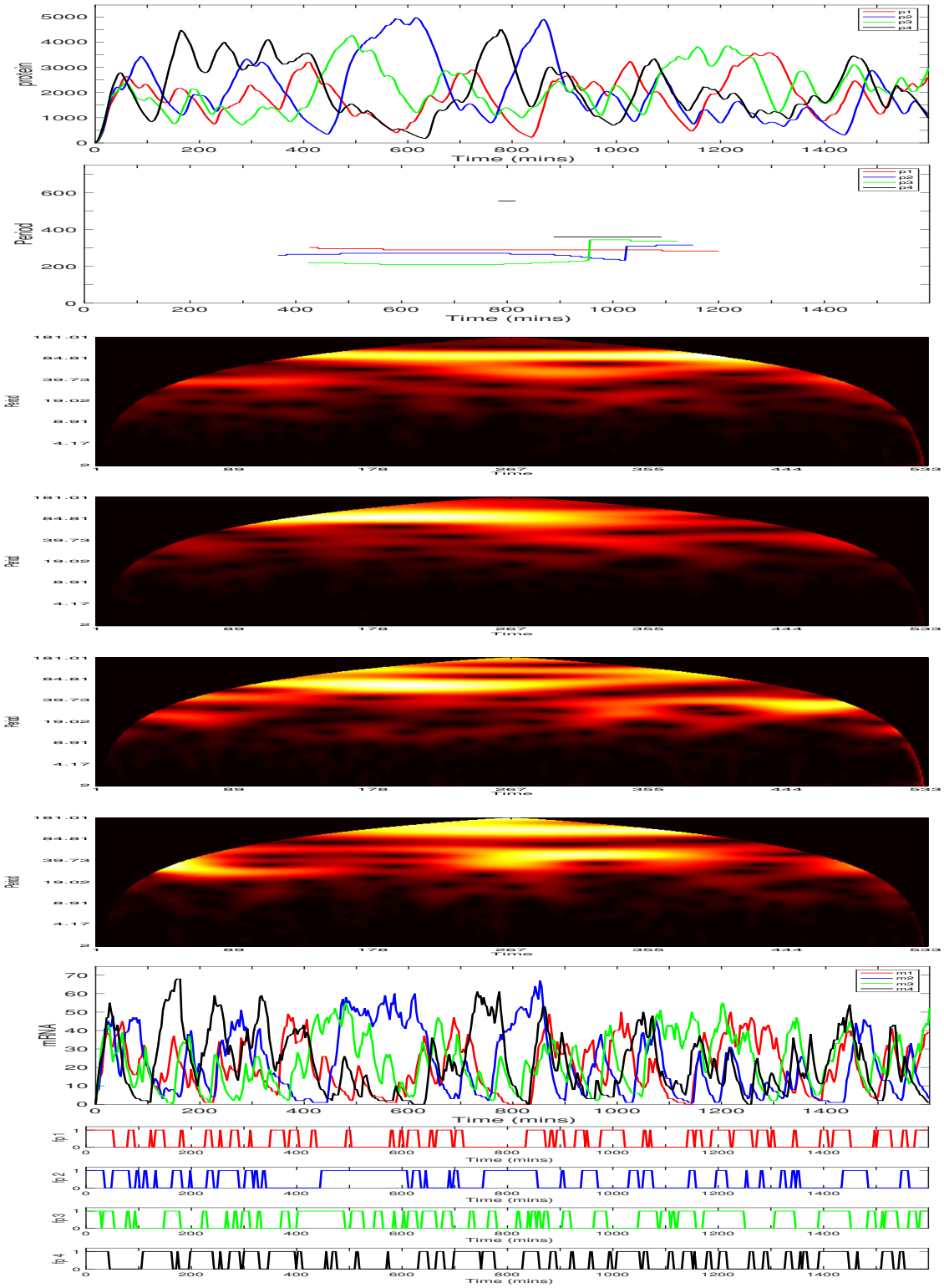

Figure 5.67: *The four gene repressilator with individual gene sites equidistant apart at the nuclear membrane:* $g_1 = (0\mu m, 2.85\mu m, 0\mu m), g_2 = (2.85\mu m, 0\mu m, 0\mu m)$, $g_3 = (0\mu m, -2.85\mu m, 0\mu m)$ *and* $g_4 = (-2.85\mu m, 0\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products. and Gene4/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

For $D = 1 \times 10^{-12} m^2 min^{-1}$ the average period for 100 trajectories was $\bar{T} = 400$, ranging between, $193 \leq \bar{T} \leq 555$. The average protein number was $\bar{p} = 1597$, ranging between $242 \leq \bar{p} \leq 2707$. The average peak value of protein number was $\bar{p}_{peak} = 3261$, ranging between $1370 \leq \bar{p}_{peak} \leq 3889$.

For $D = 1 \times 10^{-10} m^2 min^{-1}$ the average period for 100 trajectories was $\bar{T} = 300$, ranging between, $131 \leq \bar{T} \leq 555$. The average protein number was $\bar{p} = 2015$, ranging between $1222 \leq \bar{p} \leq 2942$. The average peak value of protein number was $\bar{p}_{peak} = 4497$, ranging between $2978 \leq \bar{p}_{peak} \leq 5781$.

For gene sites at opposite ends of the x-axis, close to the nucleus membrane, we see similar behaviour as per the case where gene sites are clustered at the origin. However, there is a difference in that the 'oscillations' observed for the lower diffusion regime seem slightly more defined. Comparing to the case of the two-gene and three-gene repressilator, where stronger oscillatory type activity was seen as we increased from two to three genes, we again see and increase in oscillatory type behaviour with the addition of the fourth gene. Since each gene site is situated in close proximity to the cytoplasm, their corresponding protein will reach greater levels, than the origin case. However, this time the distance from the synthesised protein and its target gene has decreased once again with the new configuration given to the positions of the genes with the addition of gene 4. For the faster diffusion regimes, oscillatory behaviour seen appears to be noisier than the slower diffusion regime. However, WAVOS, detected lower modal periods for the faster diffusion regime. In consideration of the heat maps, we can safely say that there are higher modes in the system that will not be picked up by WAVOS in the time frame chosen. Again suggesting the multi-modal behaviour of the n-gene repressilator. More investigation needs to be applied to looking into any significance in these high and low period modes to GRN dynamics.

We now consider the gene site locations to be $g_1 = (x_1, y_1, z_1) = (0\mu m, 2.85\mu m, 0\mu m)$,

$g_2 = (x_2, y_2, z_2) = (\frac{1}{\sqrt{2}}\mu m, 2.85 - \frac{1}{\sqrt{2}}\mu m, 0\mu m)$, $g_3 = (x_3, y_3, z_3) = (0\mu m, 2.85 - \sqrt{2}\mu m, 0\mu m)$ and $g_4 = (x_4, y_4, z_4) = (-\frac{1}{\sqrt{2}}\mu m, 2.85 - \frac{1}{\sqrt{2}}\mu m, 0\mu m)$ for gene 1, 2, 3 and 4 respectively i.e we translate the, origin clustered gene sites positions, approximately 2.14μ along the positive y-axis. We term this position a cluster of gene sites close to the nuclear membrane. The first set of three simulations are for the smaller diffusion regime, $D = 1 \times 10^{-12}$ and the second set of three simulations are for the faster diffusion regime, $D = 1 \times 10^{-10}$.

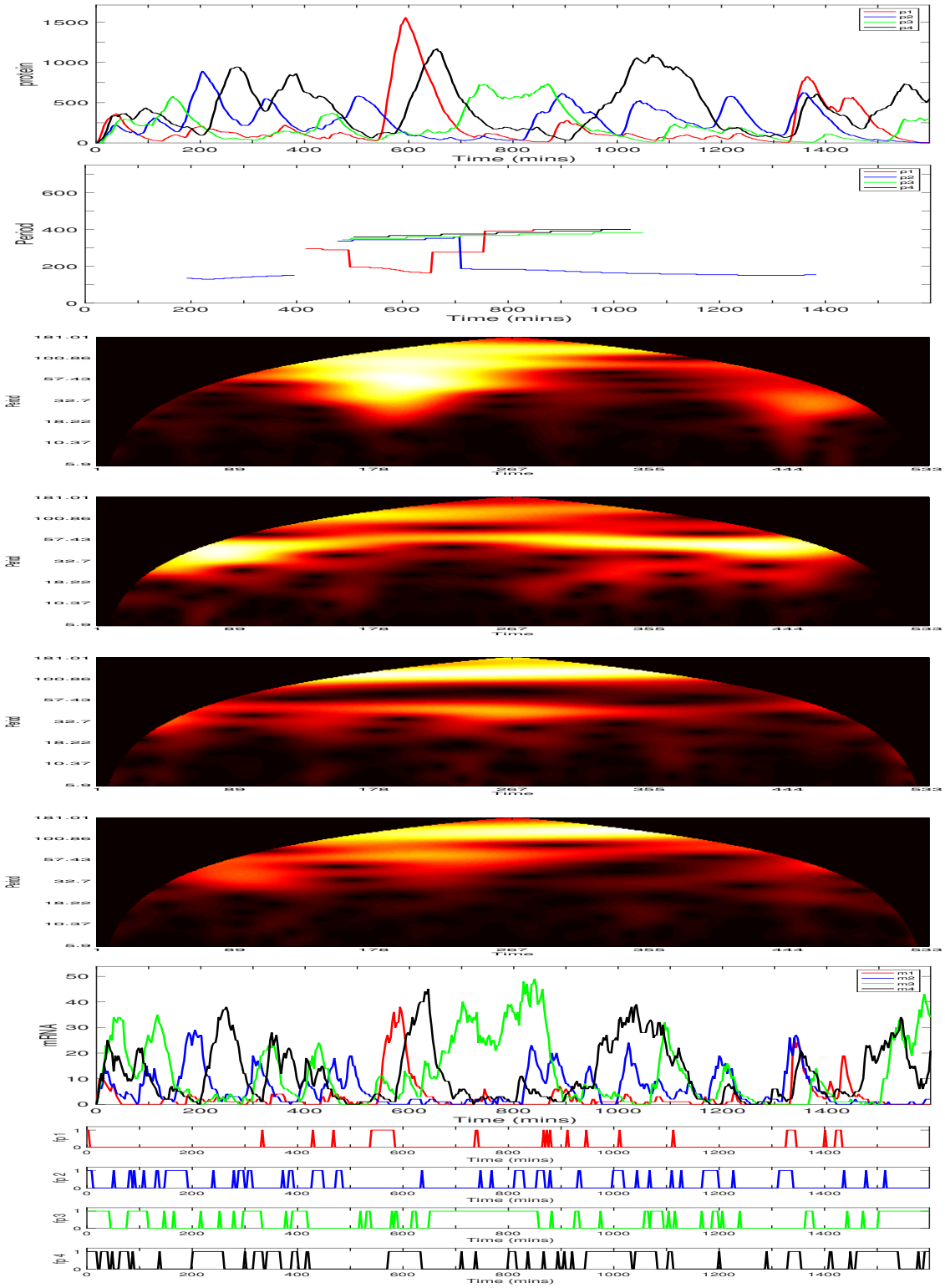

Figure 5.68: *The four gene repressilator with individual gene sites clustered together close to the nuclear membrane:* $g_1 = (x_1, y_1, z_1) = (0\mu m, 2.85\mu m, 0\mu m)$, $g_2 = (x_2, y_2, z_2) = (\frac{1}{\sqrt{2}}\mu m, 2.85 - \frac{1}{\sqrt{2}}\mu m, 0\mu m)$, $g_3 = (x_3, y_3, z_3) = (0\mu m, 2.85 - \sqrt{2}\mu m, 0\mu m)$ *and* $g_4 = (x_4, y_4, z_4) = (-\frac{1}{\sqrt{2}}\mu m, 2.85 - \frac{1}{\sqrt{2}}\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products. and Gene4/products. Top to bottom: time series of protein, Period, heat maps for periods of protein 1, 2, 3 and 4, mRNA and free promoter. For each species the total copy number was counted over the entire cell.. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

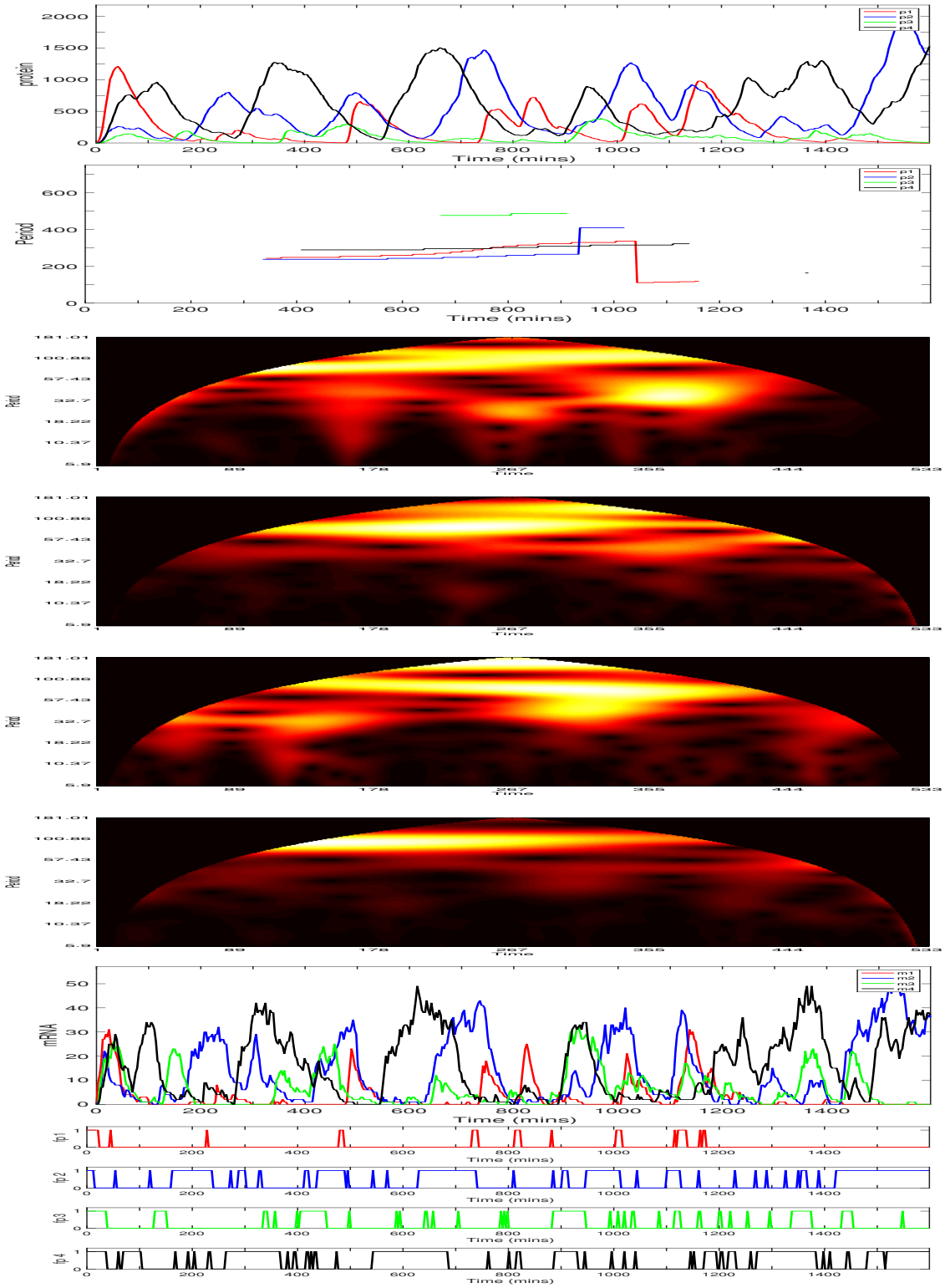

Figure 5.69: *The four gene repressilator with individual gene sites clustered together close to the nuclear membrane:* $g_1 = (x_1, y_1, z_1) = (0\mu m, 2.85\mu m, 0\mu m)$, $g_2 = (x_2, y_2, z_2) = (\frac{1}{\sqrt{2}}\mu m, 2.85 - \frac{1}{\sqrt{2}}\mu m, 0\mu m)$, $g_3 = (x_3, y_3, z_3) = (0\mu m, 2.85 - \sqrt{2}\mu m, 0\mu m)$ *and* $g_4 = (x_4, y_4, z_4) = (-\frac{1}{\sqrt{2}}\mu m, 2.85 - \frac{1}{\sqrt{2}}\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products. and Gene4/products. Top to bottom: time series of protein, Period, heat maps for periods of protein 1, 2, 3 and 4, mRNA and free promoter. For each species the total copy number was counted over the entire cell.. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

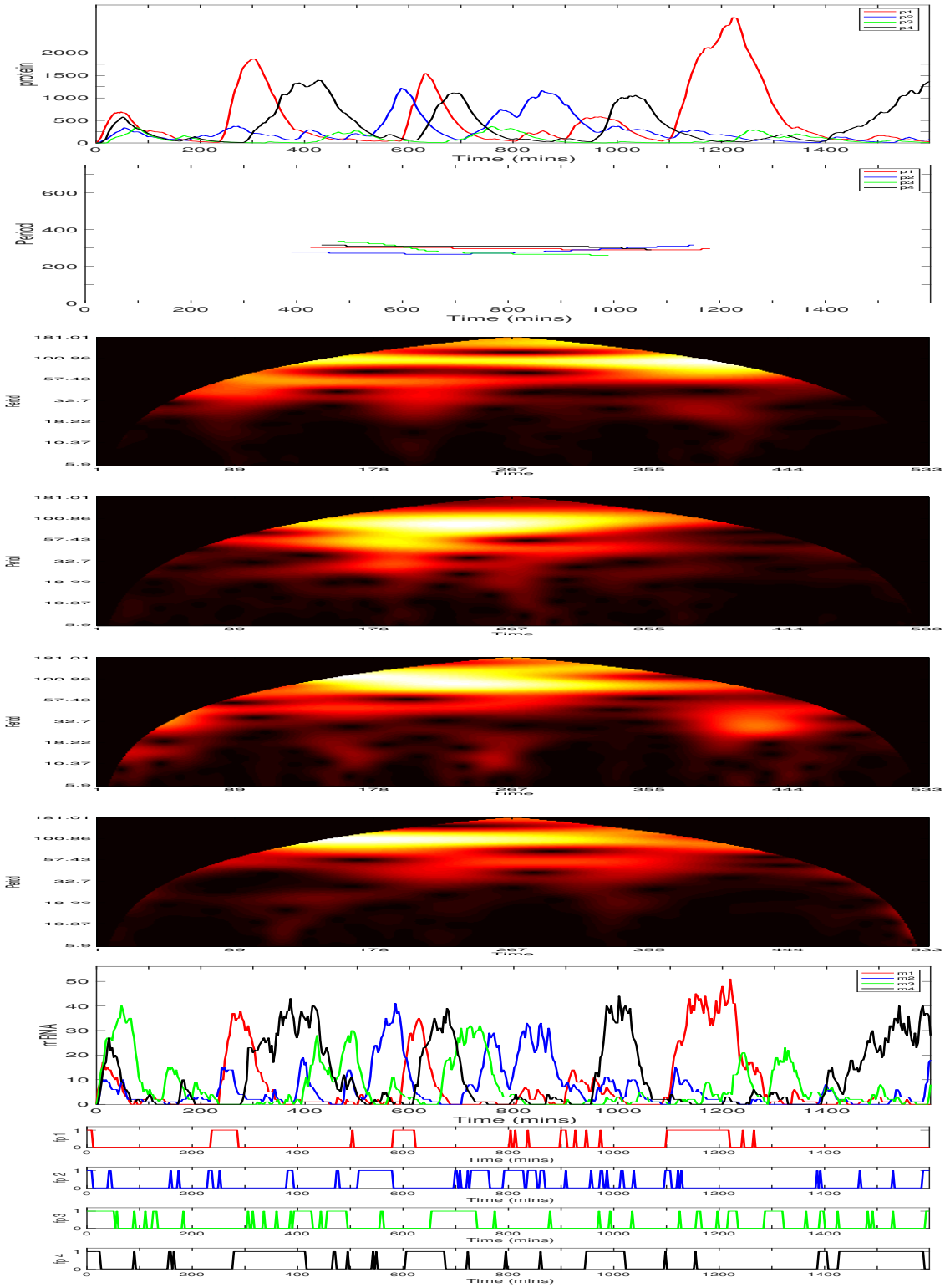

Figure 5.70: *The four gene repressilator with individual gene sites clustered together close to the nuclear membrane:* $g_1 = (x_1, y_1, z_1) = (0\mu m, 2.85\mu m, 0\mu m)$, $g_2 = (x_2, y_2, z_2) = (\frac{1}{\sqrt{2}}\mu m, 2.85 - \frac{1}{\sqrt{2}}\mu m, 0\mu m)$, $g_3 = (x_3, y_3, z_3) = (0\mu m, 2.85 - \sqrt{2}\mu m, 0\mu m)$ *and* $g_4 = (x_4, y_4, z_4) = (-\frac{1}{\sqrt{2}}\mu m, 2.85 - \frac{1}{\sqrt{2}}\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products. and Gene4/products. Top to bottom: time series of protein, Period, heat maps for periods of protein 1, 2, 3 and 4, mRNA and free promoter. For each species the total copy number was counted over the entire cell.. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

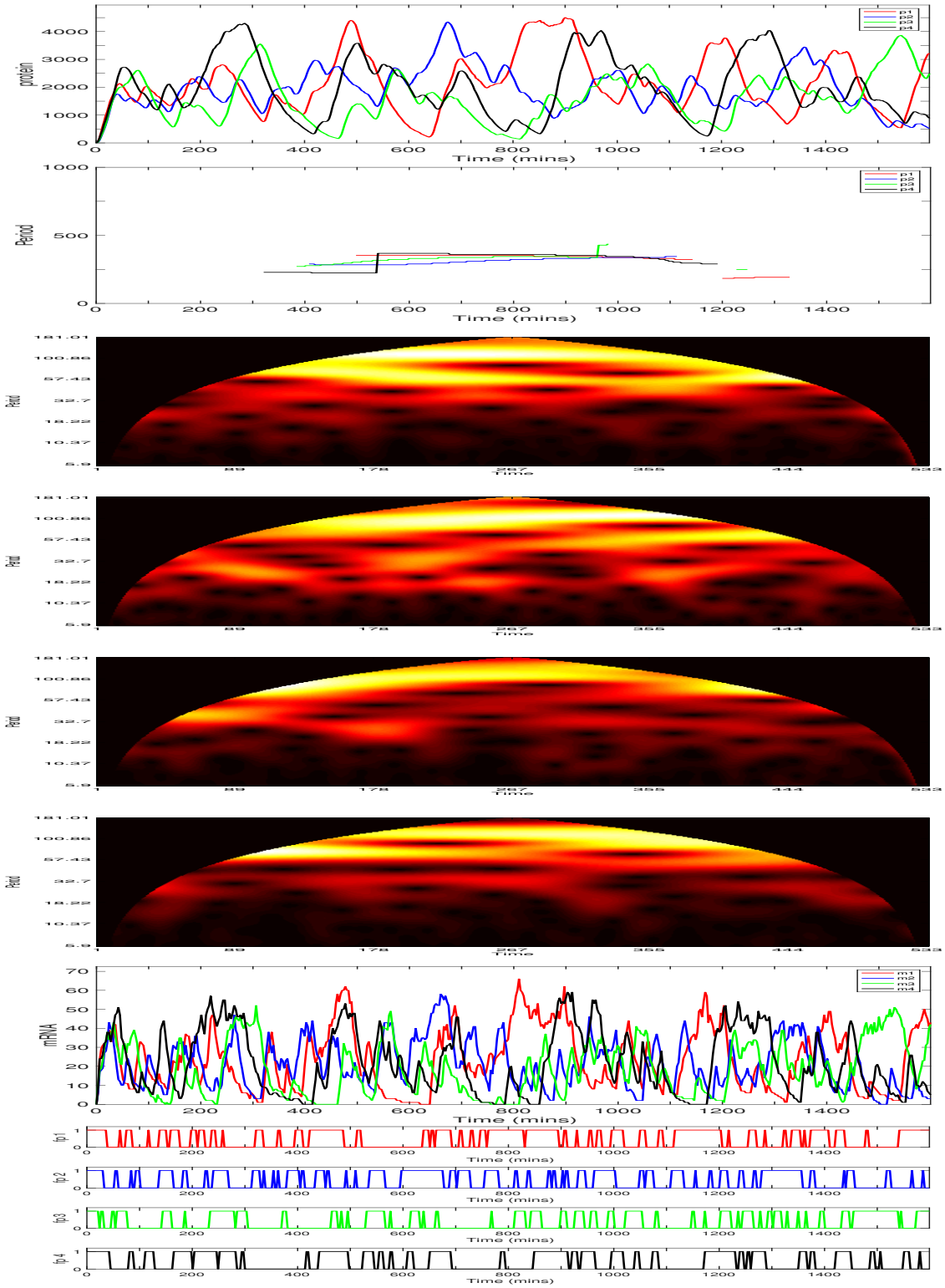

Figure 5.71: *The four gene repressilator with individual gene sites clustered together at the nuclear membrane:* $g_1 = (0\mu m, 2.1 + \frac{1}{\sqrt{2}}\mu m, 0\mu m)$, $g_2 = (\frac{1}{\sqrt{2}}\mu m, 2.1\mu m, 0\mu m)$, $g_3 = (0\mu m, 2.1 - \frac{1}{\sqrt{2}}\mu m, 0\mu m)$ *and* $g_4 = (-\frac{1}{\sqrt{2}}\mu m, 2.1\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products and Gene3/products. Top to bottom: Protein, Period, Period heat maps, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

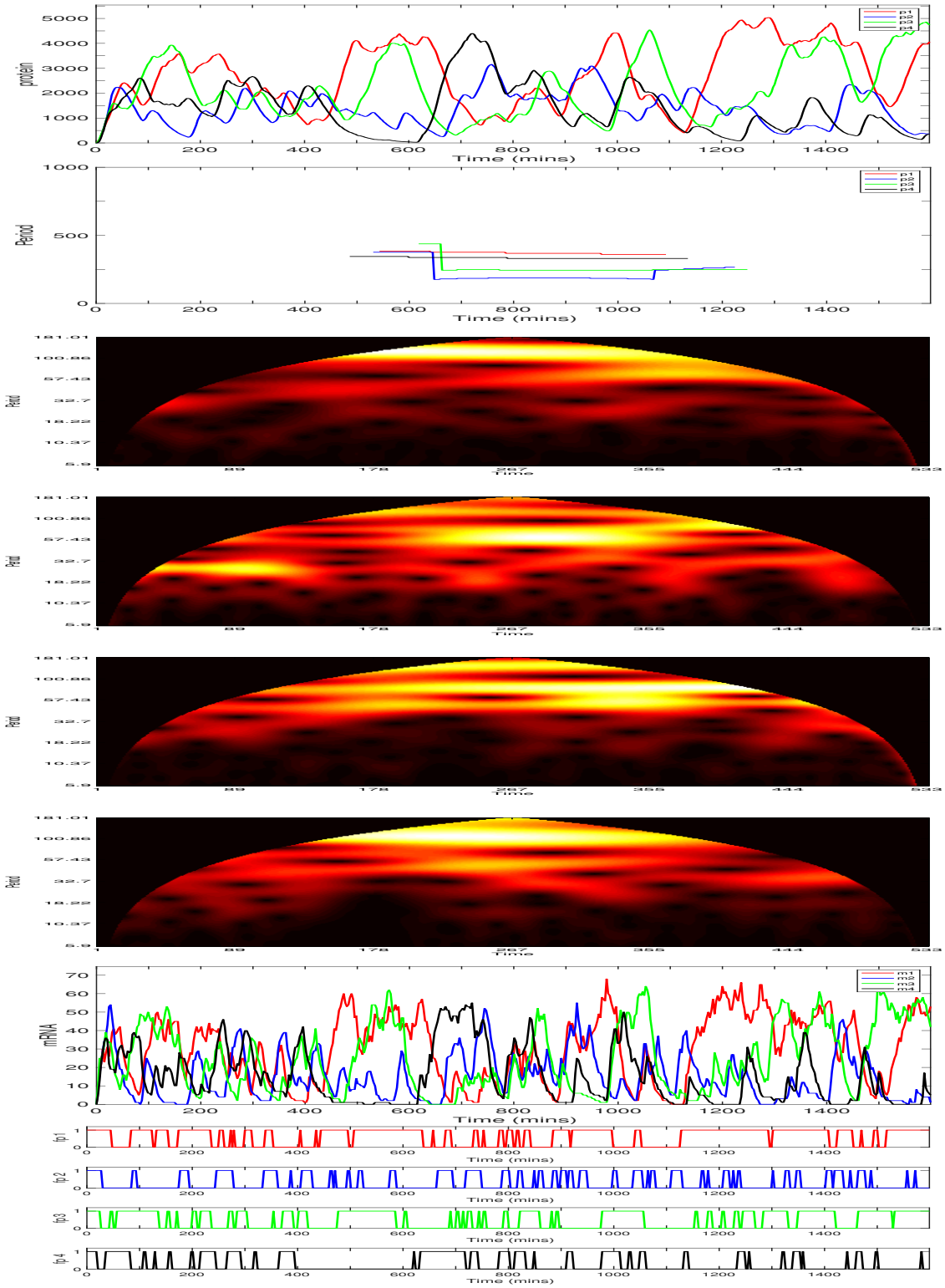

Figure 5.72: *The four gene repressilator with individual gene sites clustered together at the nuclear membrane:* $g_1 = (0\mu m, 2.1 + \frac{1}{\sqrt{2}}\mu m, 0\mu m)$, $g_2 = (\frac{1}{\sqrt{2}}\mu m, 2.1\mu m, 0\mu m)$, $g_3 = (0\mu m, 2.1 - \frac{1}{\sqrt{2}}\mu m, 0\mu m)$ *and* $g_4 = (-\frac{1}{\sqrt{2}}\mu m, 2.1\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products and Gene3/products. Top to bottom: Protein, Period, Period heat maps, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

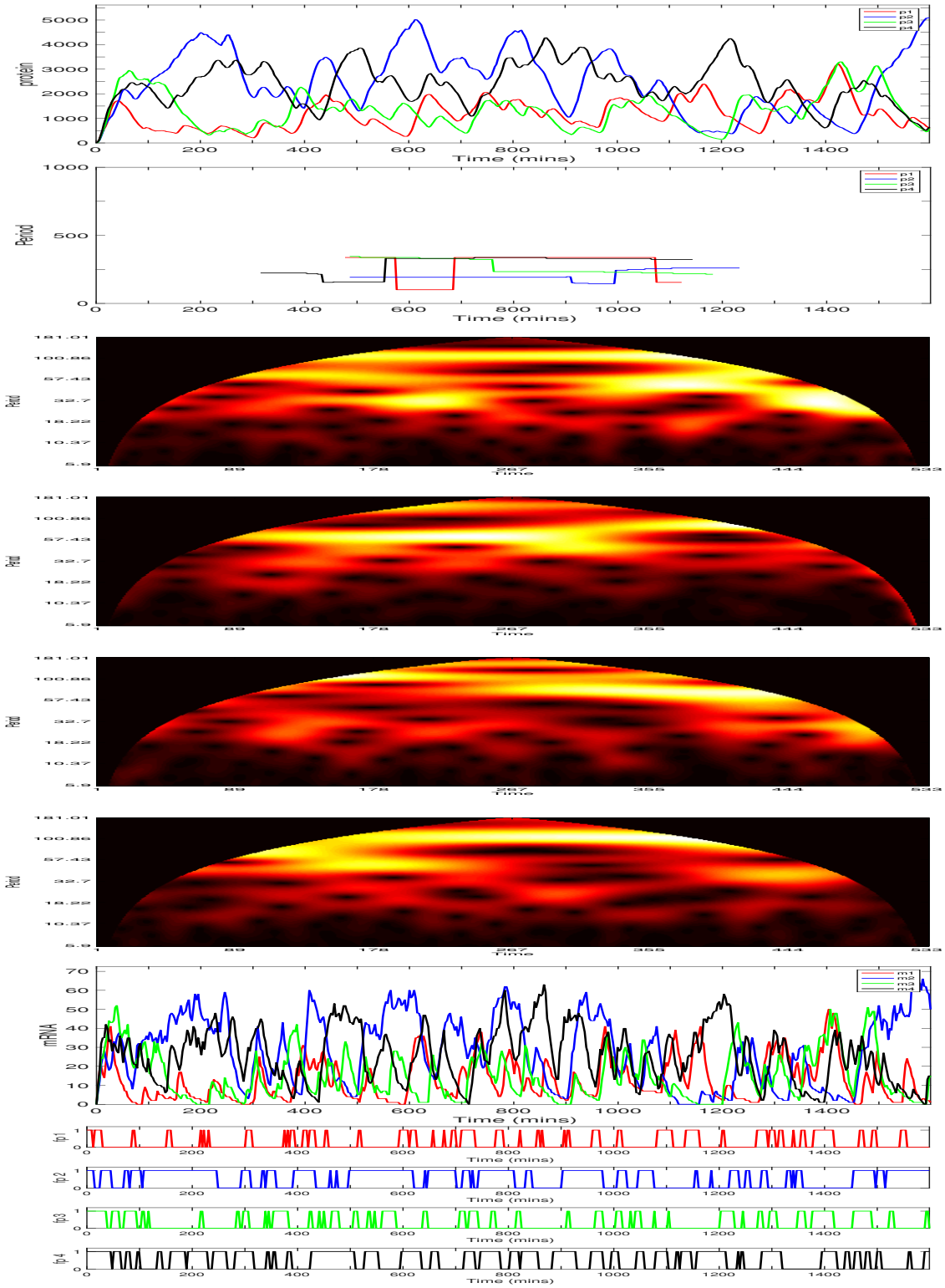

Figure 5.73: *The four gene repressilator with individual gene sites clustered together at the nuclear membrane:* $g_1 = (0\mu m, 2.1 + \frac{1}{\sqrt{2}}\mu m, 0\mu m)$, $g_2 = (\frac{1}{\sqrt{2}}\mu m, 2.1\mu m, 0\mu m)$, $g_3 = (0\mu m, 2.1 - \frac{1}{\sqrt{2}}\mu m, 0\mu m)$ *and* $g_4 = (-\frac{1}{\sqrt{2}}\mu m, 2.1\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products and Gene3/products. Top to bottom: Protein, Period, Period heat maps, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

For $D = 1 \times 10^{-12} m^2 min^{-1}$ the average period for 100 trajectories was $\bar{T} = 350$, ranging between, $181 \leq \bar{T} \leq 555$. The average protein number was $\bar{p} = 278$, ranging between $56 \leq \bar{p} \leq 745$. The average peak value of protein number was $\bar{p}_{peak} = 1848$, ranging between $280 \leq \bar{p}_{peak} \leq 3508$.

For $D = 1 \times 10^{-10} m^2 min^{-1}$ the average period for 100 trajectories was $\bar{T} = 304$, ranging between, $109 \leq \bar{T} \leq 555$. The average protein number was $\bar{p} = 1894$, ranging between $1234 \leq \bar{p} \leq 2621$. The average peak value of protein number was $\bar{p}_{peak} = 1874$, ranging between $1166 \leq \bar{p}_{peak} \leq 2814$.

WAVOS calculates periods that are between 5 and 10 hours long. However, we remain cautious of the periods detected by WAVOS, again speculating that these are lower modes in the system. The oscillatory behaviour observed still remains elusive and in consideration of the heat maps, we speculate that their will be higher modal periods in the system.

For the slower diffusion regime we see that the proximity of the gene sites to the cytoplasm has a great effect on the levels of protein. Protein one having the greater advantage as it is closest and second, protein two and four. The protein that forms the least amount is protein three, due to the fact they are furthest away from the cytoplasm. For the faster diffusion regime, there is no bias, as the increase in diffusion creates the affect that all gene sites are close the cytoplasm.

See 5.74 to see a summarising plot of protein copy number against the diffusion parameter. Generally, as diffusion is increased, the copy number of protein increases. In terms of gene site configurations, as the gene sites are moved closer to the nucleus membrane, the copy number increased, specifically for the slower diffusion regime. However, there is no significant difference in copy number due to gene site configuration for the faster diffusion regime.

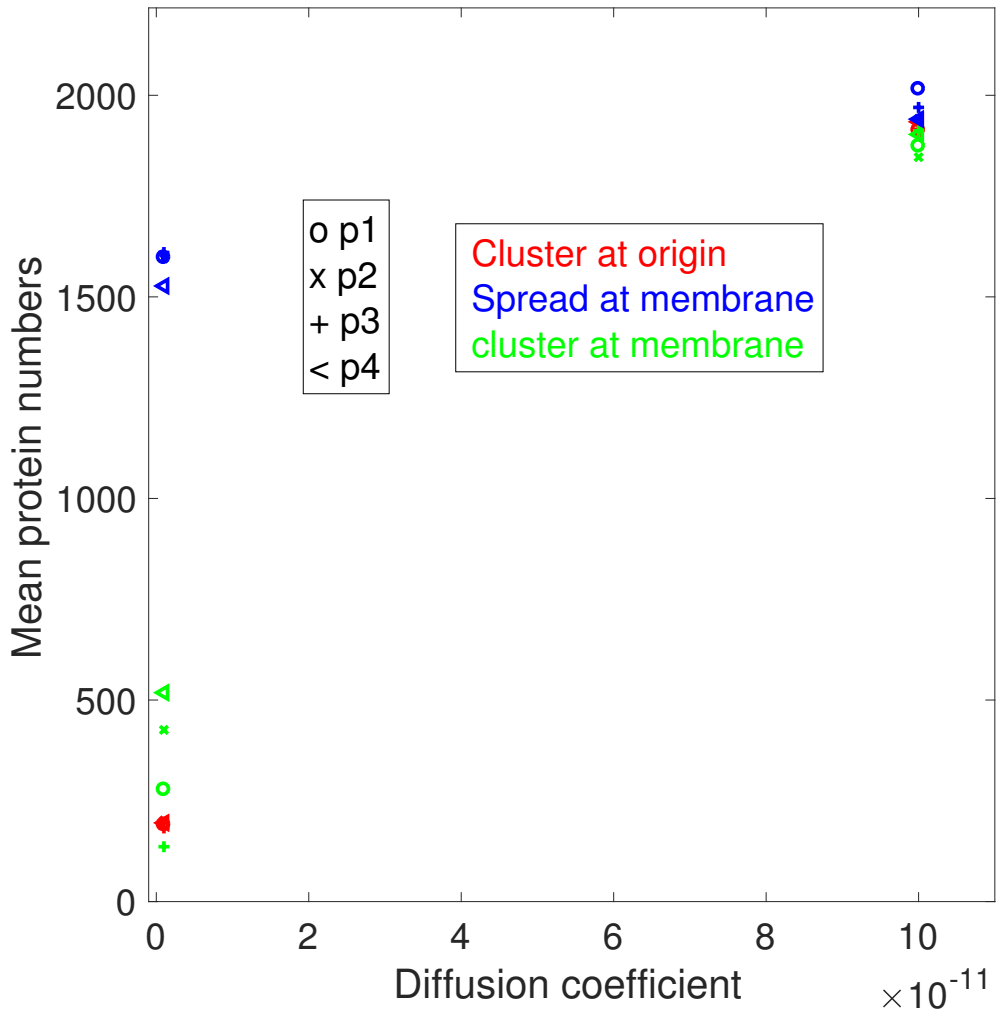

Figure 5.74: *Left: The mean value of protein calculated for the four-gene repressilator against the diffusion parameter. Right: The mean period calculated for the four-gene repressilator agains the diffusion parameter. Red denotes the gene site cluster at the origin, blue denotes the gene sites at the membrane on opposite ends of the x-axis and green denotes the cluster of genes close to the nuclear membrane.*

5.5 The five gene repressilator

Now we consider the five gene repressilator with individual gene sites clustered together at the origin. Each gene site lying on a vertex of a pentagon with sides $1\mu m$ long. $(x_1, y_1, z_1) = (0\mu m, gx\mu m, 0\mu m)$, $(x_2, y_2, z_3) = (gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, $(x_3, y_3, z_3) = (gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10)\mu m, 0\mu m)$, $(x_4, y_4, z_4) = (-gx\cos(3\pi/10)\mu m, -gx\s$ and $(x_5, y_5, z_5) = (-gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, for gene 1, 2, 3, 4 and 5 respectively. Where $gx = 1/2\sin(\pi/5)$.

The first set of three simulations are for the smaller diffusion regime, $D = 1 \times 10^{-12}$ and the second set of three simulations are for the faster diffusion regime, $D = 1 \times 10^{-10}$.

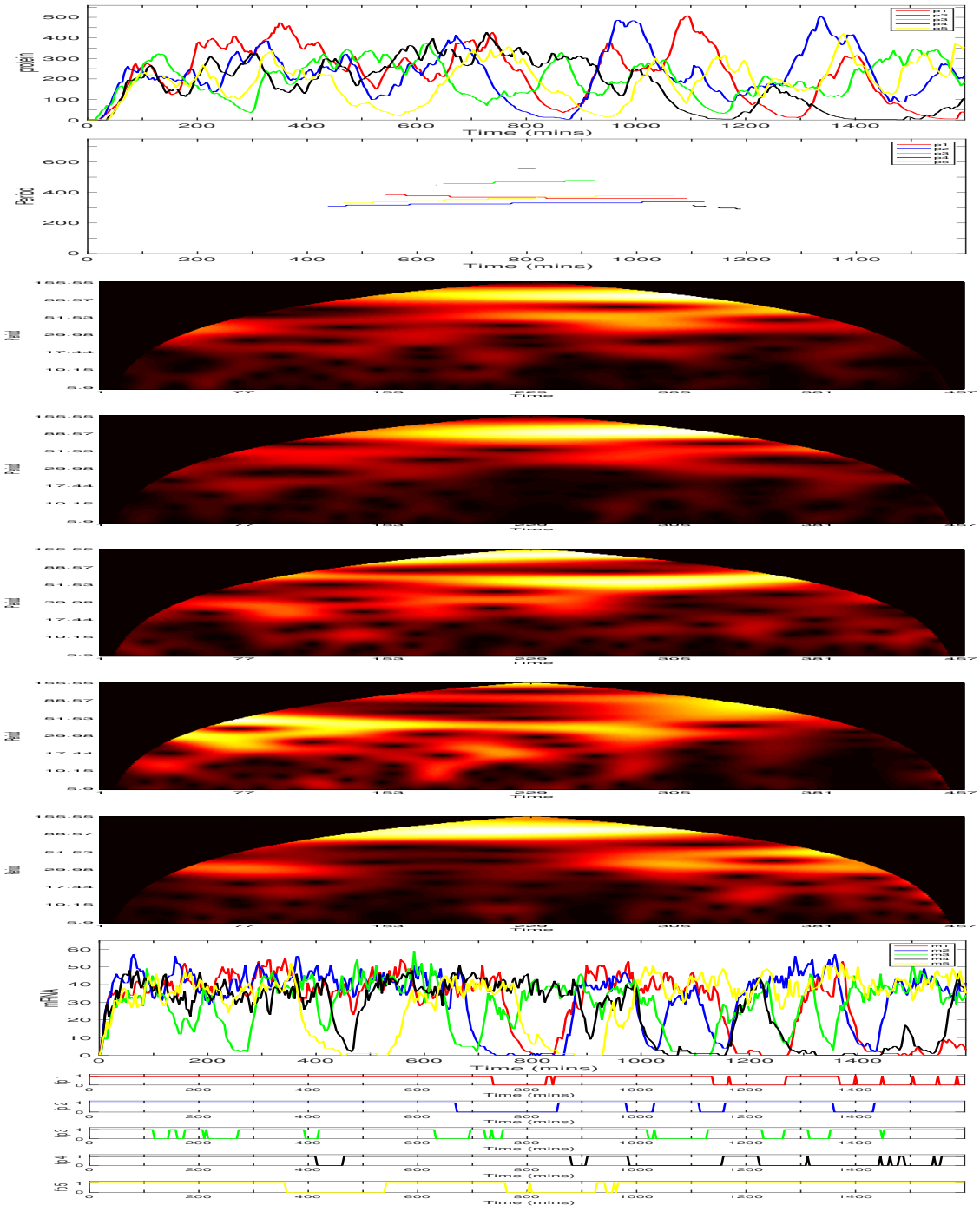

Figure 5.75: *The five gene repressilator with individual gene sites clustered together at the origin. Each gene site lying on a vertex of a pentagon with sides* $1\mu m$ *long.* $g_1 = (0\mu m, gx\mu m, 0\mu m)$, $g_2 = (gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, $g_3 = (gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10)\mu m, 0\mu m)$, $g_4 = (-gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10)\mu m, 0\mu m)$ *and* $g_5 = (-gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, *where* $gx = 1/2\sin(\pi/5)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products and Gene5/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

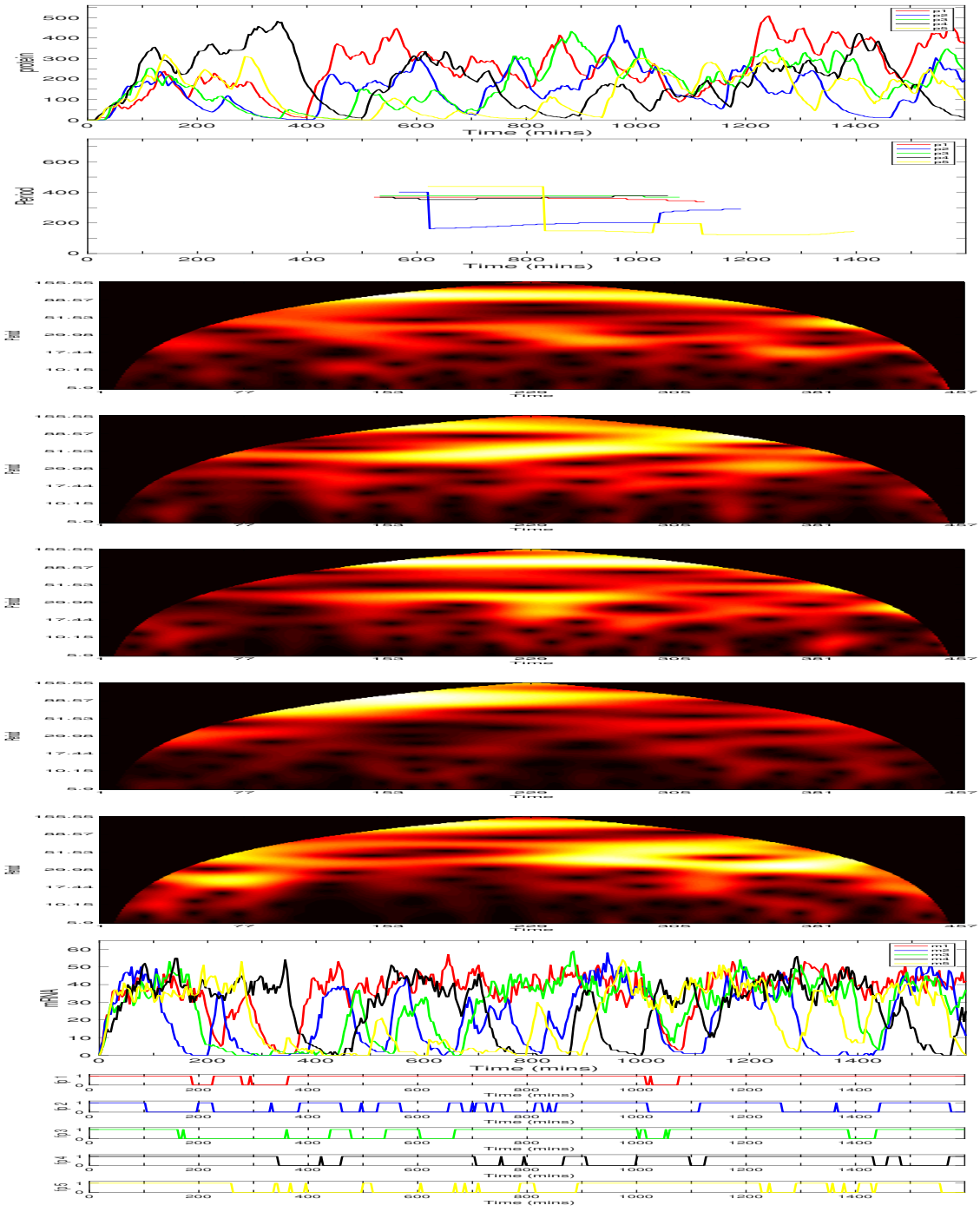

Figure 5.76: *The five gene repressilator with individual gene sites clustered together at the origin. Each gene site lying on a vertex of a pentagon with sides* $1\mu m$ *long.* $g_1 = (0\mu m, gx\mu m, 0\mu m)$, $g_2 = (gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, $g_3 = (gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10)\mu m, 0\mu m)$, $g_4 = (-gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10)\mu m, 0\mu m)$ *and* $g_5 = (-gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, *where* $gx = 1/2\sin(\pi/5)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products and Gene5/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

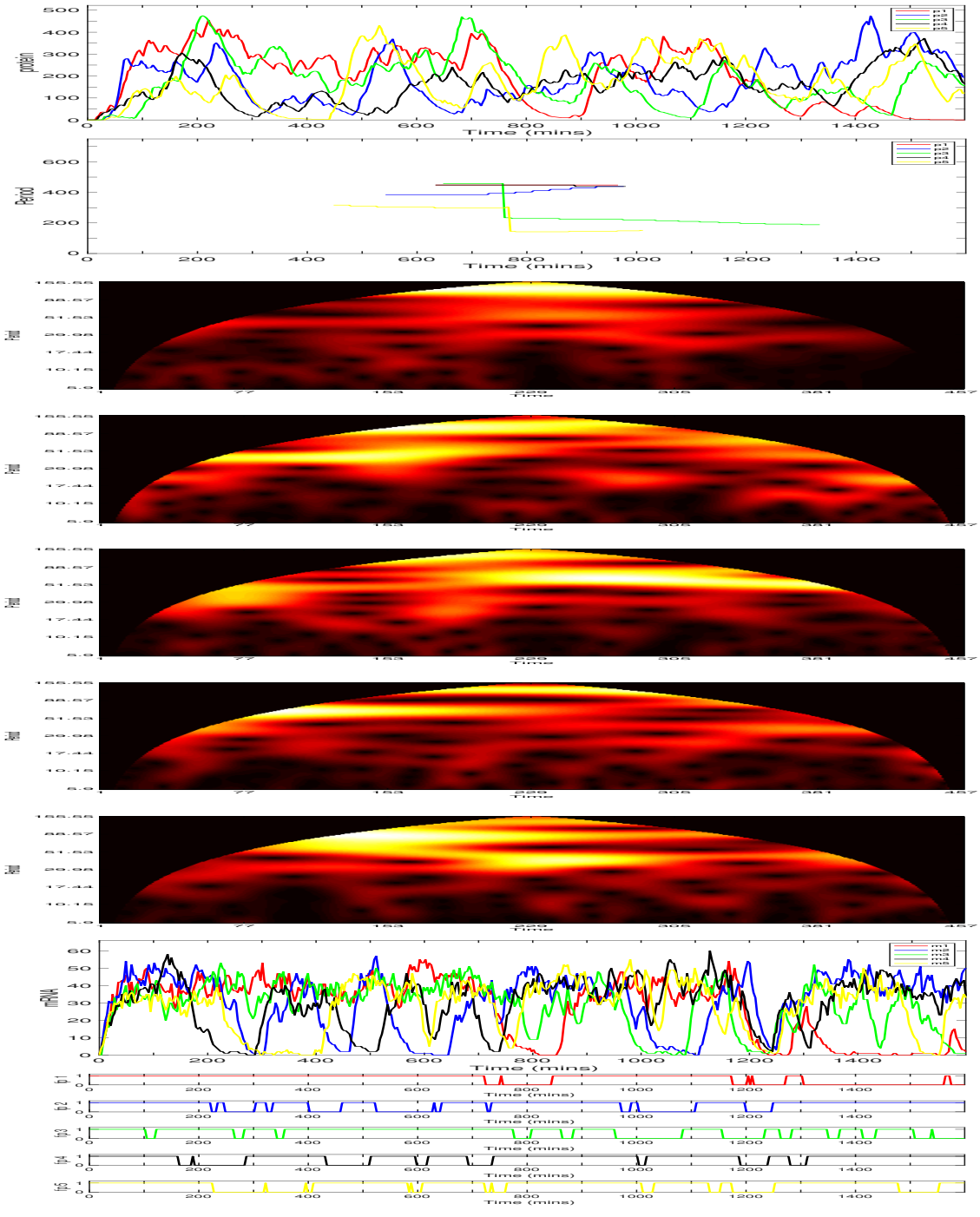

Figure 5.77: *The five gene repressilator with individual gene sites clustered together at the origin. Each gene site lying on a vertex of a pentagon with sides* $1\mu m$ *long.* $g_1 = (0\mu m, gx\mu m, 0\mu m)$, $g_2 = (gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, $g_3 = (gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10)\mu m, 0\mu m)$, $g_4 = (-gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10)\mu m, 0\mu m)$ *and* $g_5 = (-gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, *where* $gx = 1/2\sin(\pi/5)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products and Gene5/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

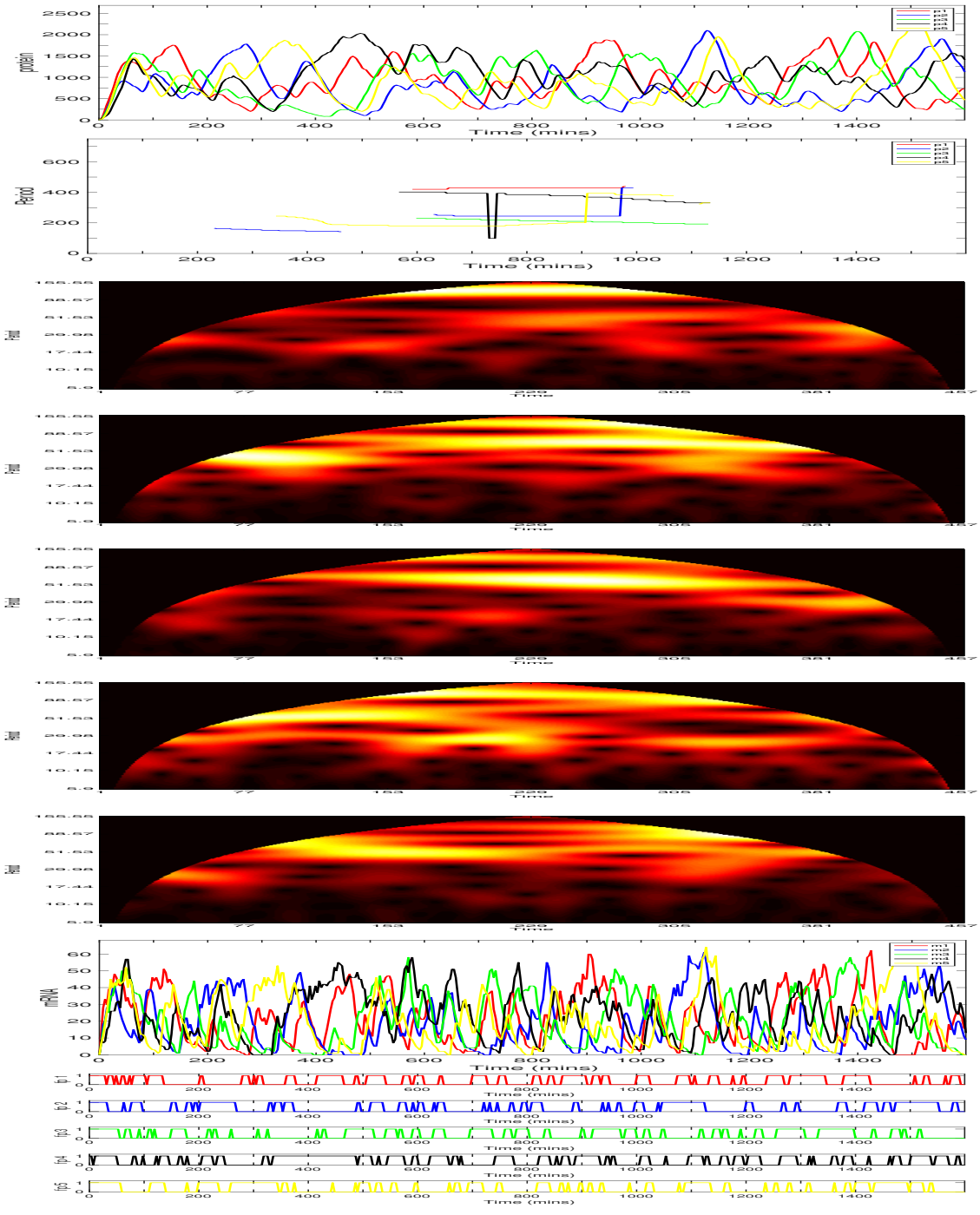

Figure 5.78: *The five gene repressilator with individual gene sites clustered together at the origin. Each gene site lying on a vertex of a pentagon with sides* $1\mu m$ *long.* $g_1 = (0\mu m, gx\mu m, 0\mu m)$, $g_2 = (gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, $g_3 = (gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10)\mu m, 0\mu m)$, $g_4 = (-gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10)\mu m, 0\mu m)$ *and* $g_5 = (-gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, *where* $gx = 1/2\sin(\pi/5)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products and Gene5/products. Top to bottom: Protein, period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

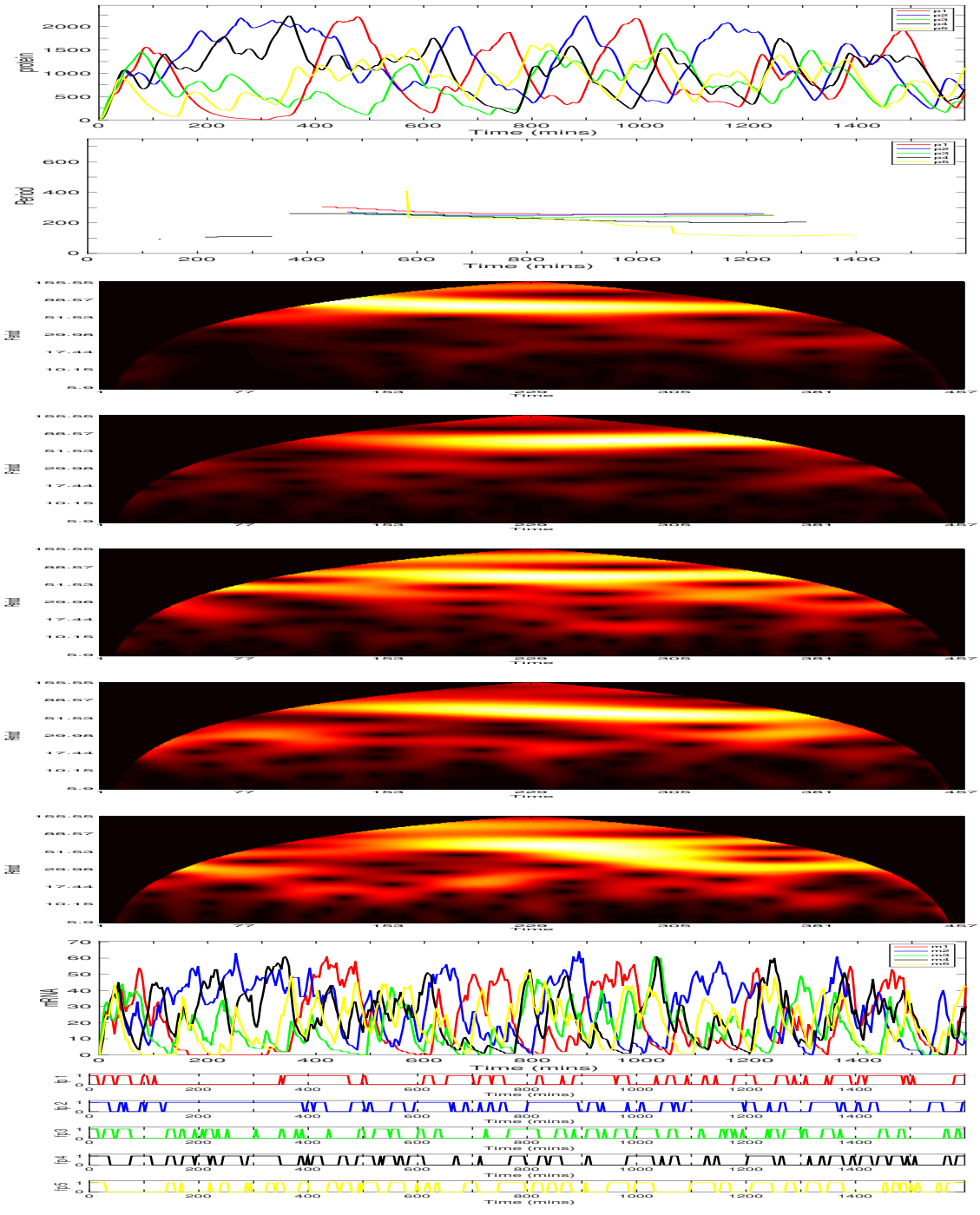

Figure 5.79: *The five gene repressilator with individual gene sites clustered together at the origin. Each gene site lying on a vertex of a pentagon with sides* $1\mu m$ *long.* $g_1 = (0\mu m, gx\mu m, 0\mu m)$, $g_2 = (gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, $g_3 = (gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10)\mu m, 0\mu m)$, $g_4 = (-gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10)\mu m, 0\mu m)$ *and* $g_5 = (-gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, *where* $gx = 1/2\sin(\pi/5)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products and Gene5/products. Top to bottom: Protein, period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

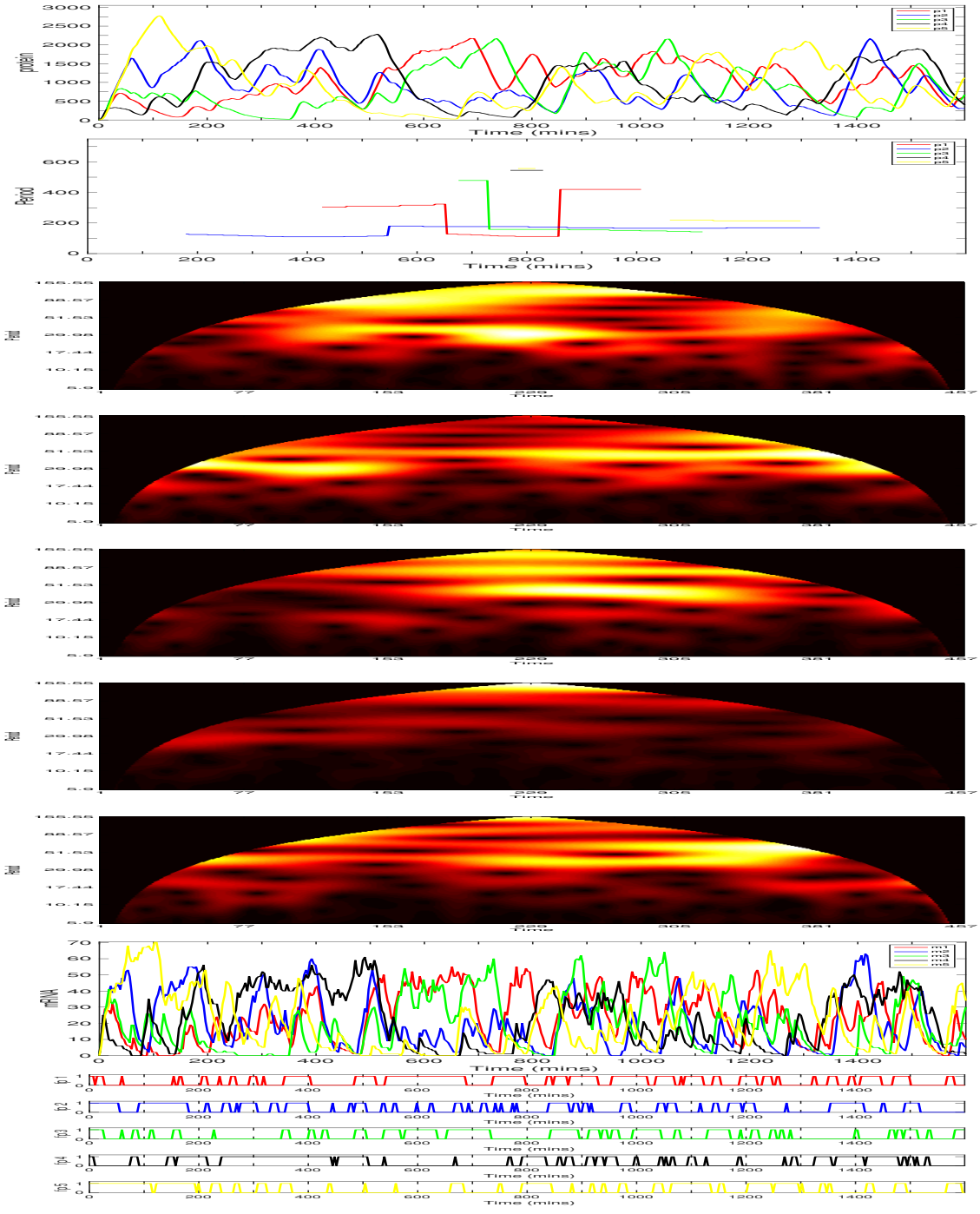

Figure 5.80: *The five gene repressilator with individual gene sites clustered together at the origin. Each gene site lying on a vertex of a pentagon with sides* $1\mu m$ *long.* $g_1 = (0\mu m, gx\mu m, 0\mu m)$, $g_2 = (gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, $g_3 = (gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10)\mu m, 0\mu m)$, $g_4 = (-gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10)\mu m, 0\mu m)$ *and* $g_5 = (-gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, *where* $gx = 1/2\sin(\pi/5)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products and Gene5/products. Top to bottom: Protein, period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

For $D = 1 \times 10^{-12} m^2 min^{-1}$ the average period for 100 trajectories was $\bar{T} = 376$, ranging between, $229 \leq \bar{T} \leq 557$. The average protein number was $\bar{p} = 178$, ranging between $94 \leq \bar{p} \leq 264$. The average peak value of protein number was $\bar{p}_{peak} = 467$, ranging between $364 \leq \bar{p}_{peak} \leq 624$.

For $D = 1 \times 10^{-10} m^2 min^{-1}$ the average period for 100 trajectories was $\bar{T} = 302$, ranging between, $132 \leq \bar{T} \leq 557$. The average protein number was $\bar{p} = 1914$, ranging between $1230 \leq \bar{p} \leq 2715$. The average peak value of protein number was $\bar{p}_{peak} = 4250$, ranging between $3341 \leq \bar{p}_{peak} \leq 5238$.

Now we consider the five-gene repressilator with individual gene sites clustered together at the origin. Each gene site lying on a vertex of a pentagon with sides $1\mu m$ long. $g_1 = (x_1, y_1, z_1) = (0\mu m, gx + 2\mu m, 0\mu m)$, $(x_2, y_2, z_3) = (gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu +$ $2m, 0\mu m)$, $(x_3, y_3, z_3) = (gx\cos(3\pi/10)\mu m, 2 - gx\sin(3\pi/10)\mu m, 0\mu m)$, $(x_4, y_4, z_4) =$ $(-gx\cos(3\pi/10)\mu m, 2 - gx\sin(3\pi/10)\mu m, 0\mu m)$ and $(x_5, y_5, z_5) = (-gx\cos(\pi/10)\mu m, gx\sin(\pi/10)$ $2\mu m, 0\mu m)$,(here $gx = 1/\sin(\pi/5)$), for gene 1, 2, 3, 4 and 5 respectively i.e. we have translated the origin clustered gene sites along the positive y-axis by $2\mu m$.

The first set of three simulations are for the smaller diffusion regime, $D = 1 \times 10^{-12}$ and the second set of three simulations are for the faster diffusion regime, $D = 1 \times 10^{-10}$.

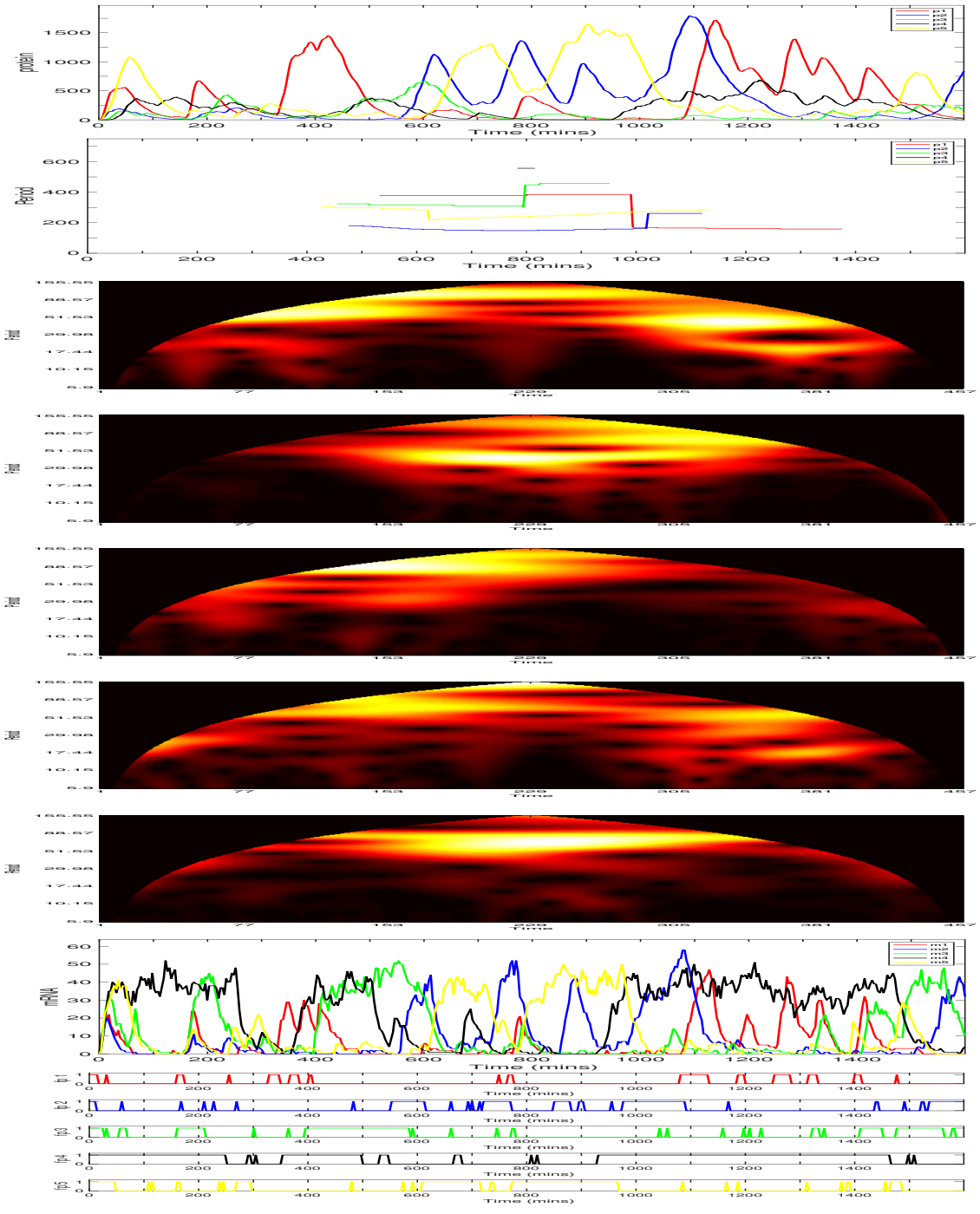

Figure 5.81: *The five gene repressilator with individual gene sites clustered together at the nuclear membrane. Each gene site lying on a vertex of a pentagon with sides* $1\mu m$ *long.* $g_1 = (0\mu m, gx + 2\mu m, 0\mu m)$, $g_2 = (gx\cos(\pi/10)\mu m, gx\sin(\pi/10) + 2\mu m, 0\mu m)$, $g_3 = (gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10) + 2\mu m, 0\mu m)$, $g_4 = (-gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10) + 2\mu m, 0\mu m)$ *and* $g_5 = (-gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, *where* $gx = 1/2\sin(\pi/5)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products and Gene5/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

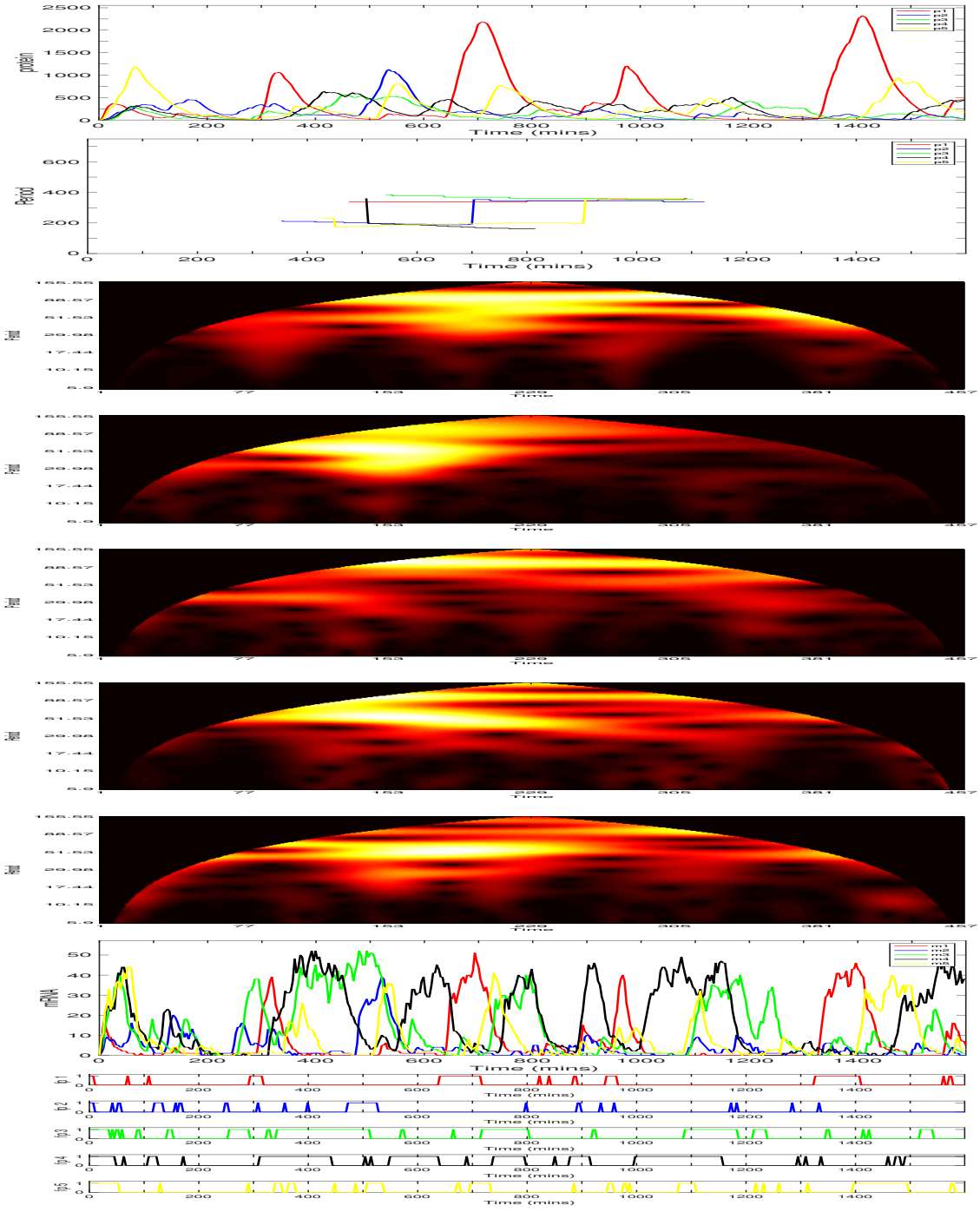

Figure 5.82: *The five gene repressilator with individual gene sites clustered together at the nuclear membrane. Each gene site lying on a vertex of a pentagon with sides* $1\mu m$ *long.* $g_1 = (0\mu m, gx + 2\mu m, 0\mu m)$, $g_2 = (gx\cos(\pi/10)\mu m, gx\sin(\pi/10) + 2\mu m, 0\mu m)$, $g_3 = (gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10) + 2\mu m, 0\mu m)$, $g_4 = (-gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10) + 2\mu m, 0\mu m)$ *and* $g_5 = (-gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, *where* $gx = 1/2\sin(\pi/5)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products and Gene5/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

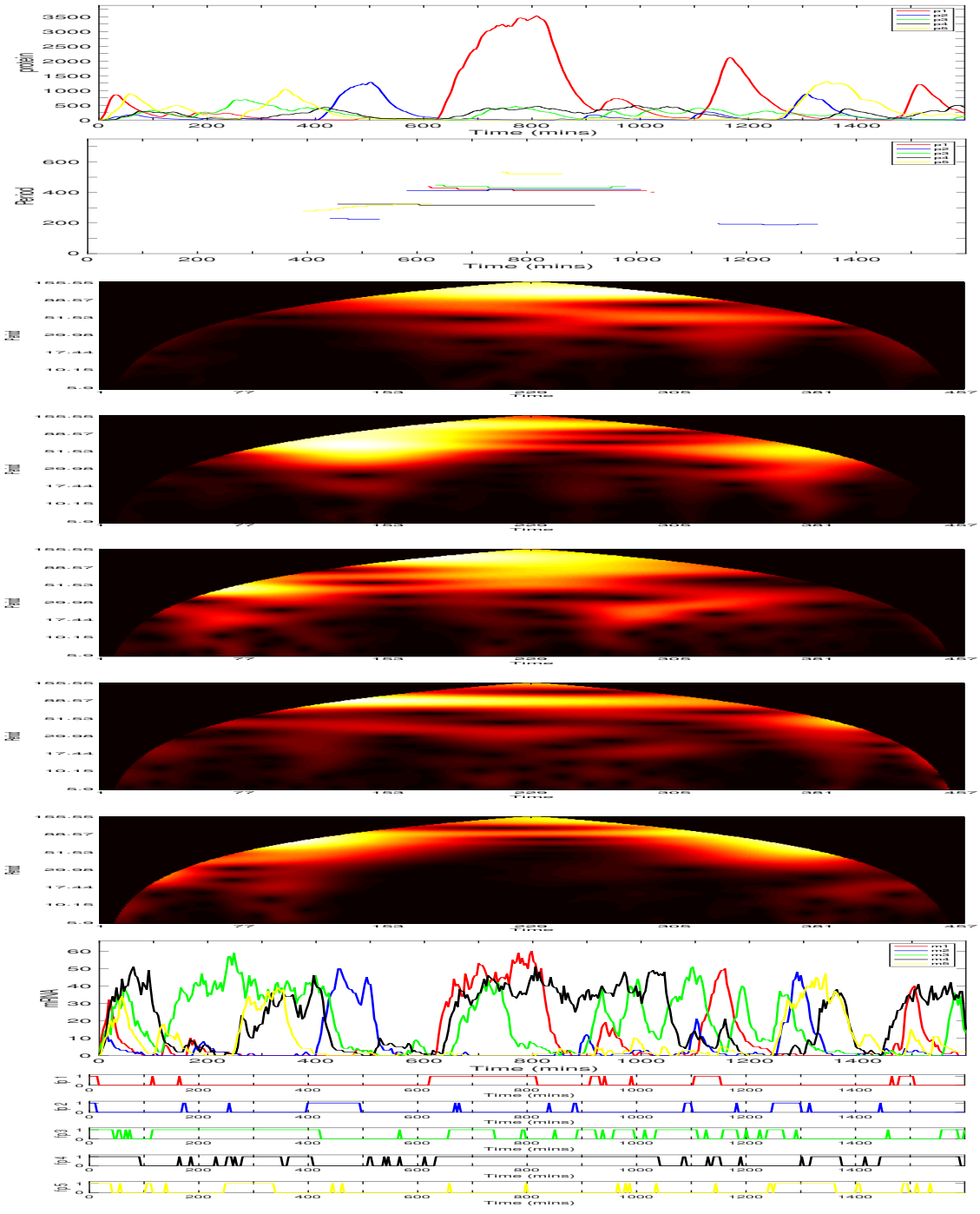

Figure 5.83: *The five gene repressilator with individual gene sites clustered together at the nuclear membrane. Each gene site lying on a vertex of a pentagon with sides* $1\mu m$ *long.* $g_1 = (0\mu m, gx + 2\mu m, 0\mu m)$, $g_2 = (gx\cos(\pi/10)\mu m, gx\sin(\pi/10) + 2\mu m, 0\mu m)$, $g_3 = (gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10) + 2\mu m, 0\mu m)$, $g_4 = (-gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10) + 2\mu m, 0\mu m)$ *and* $g_5 = (-gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, *where* $gx = 1/2\sin(\pi/5)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products and Gene5/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

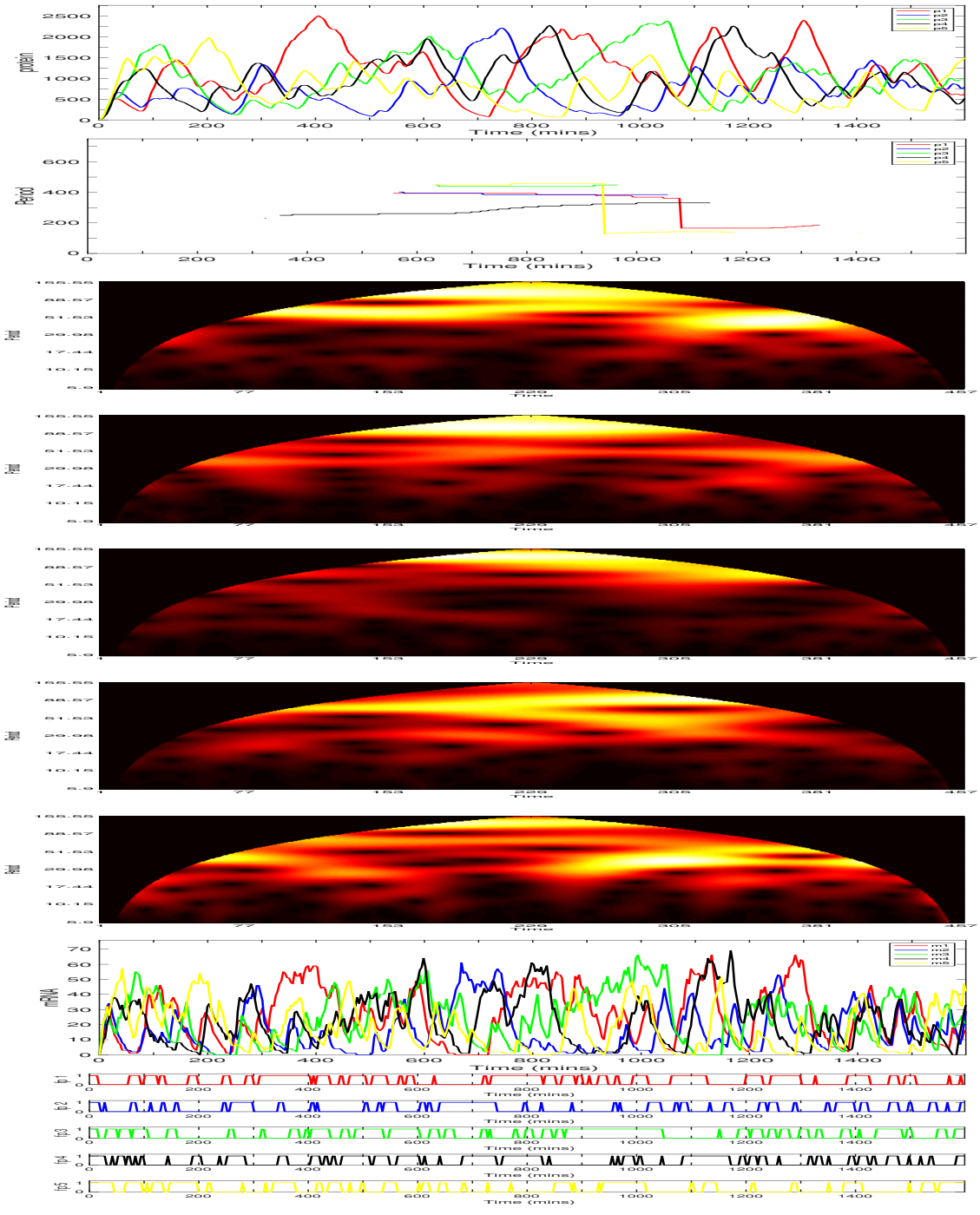

Figure 5.84: *The five gene repressilator with individual gene sites clustered together at the nuclear membrane. Each gene site lying on a vertex of a pentagon with sides* $1\mu m$ *long.* $g_1 = (0\mu m, gx + 2\mu m, 0\mu m)$, $g_2 = (gx\cos(\pi/10)\mu m, gx\sin(\pi/10) + 2\mu m, 0\mu m)$, $g_3 = (gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10) + 2\mu m, 0\mu m)$, $g_4 = (-gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10) + 2\mu m, 0\mu m)$ *and* $g_5 = (-gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, *where* $gx = 1/2\sin(\pi/5)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products and Gene5/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

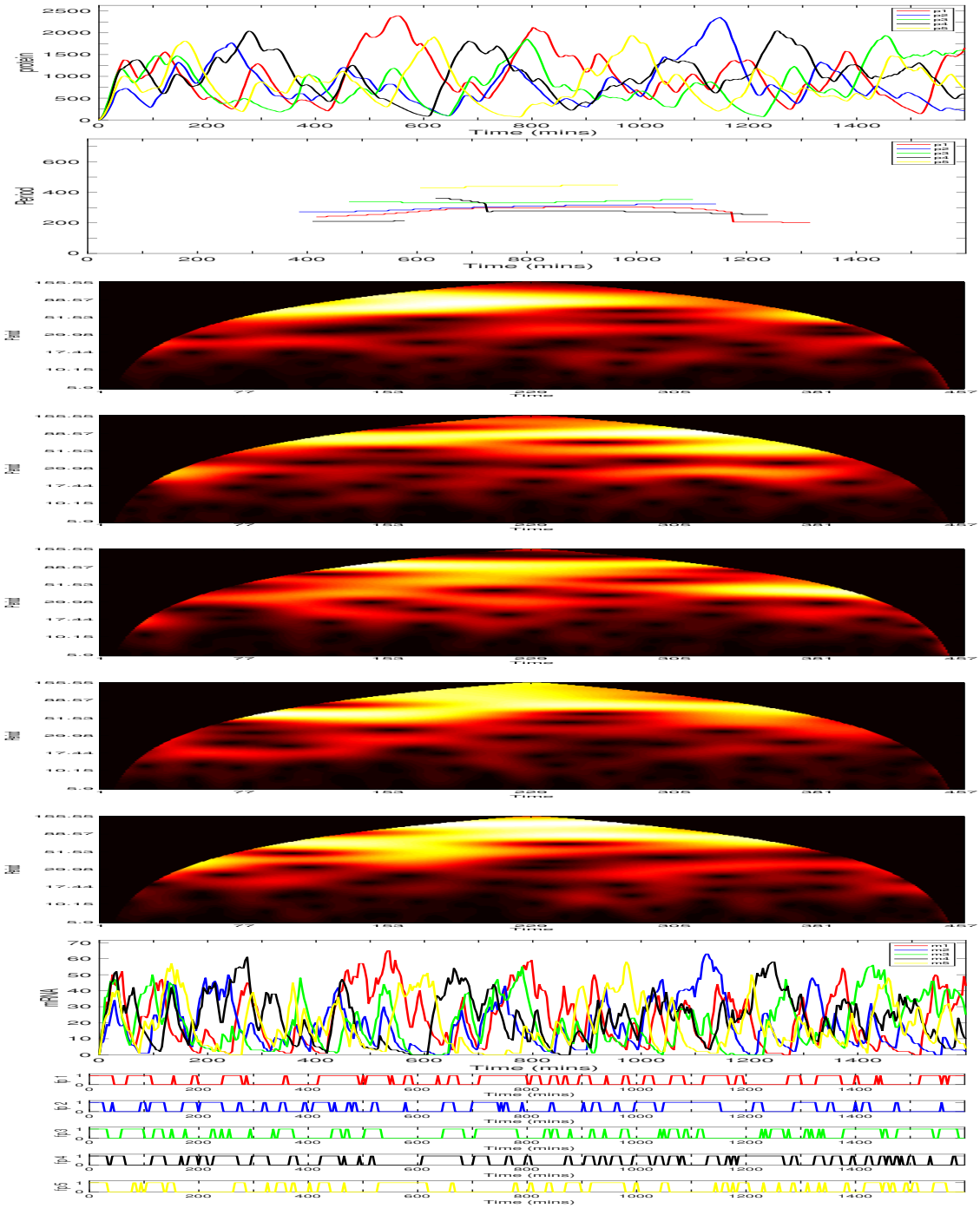

Figure 5.85: *The five gene repressilator with individual gene sites clustered together at the nuclear membrane. Each gene site lying on a vertex of a pentagon with sides* $1\mu m$ *long.* $g_1 = (0\mu m, gx + 2\mu m, 0\mu m)$, $g_2 = (gx\cos(\pi/10)\mu m, gx\sin(\pi/10) + 2\mu m, 0\mu m)$, $g_3 = (gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10) + 2\mu m, 0\mu m)$, $g_4 = (-gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10) + 2\mu m, 0\mu m)$ *and* $g_5 = (-gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, *where* $gx = 1/2\sin(\pi/5)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products and Gene5/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

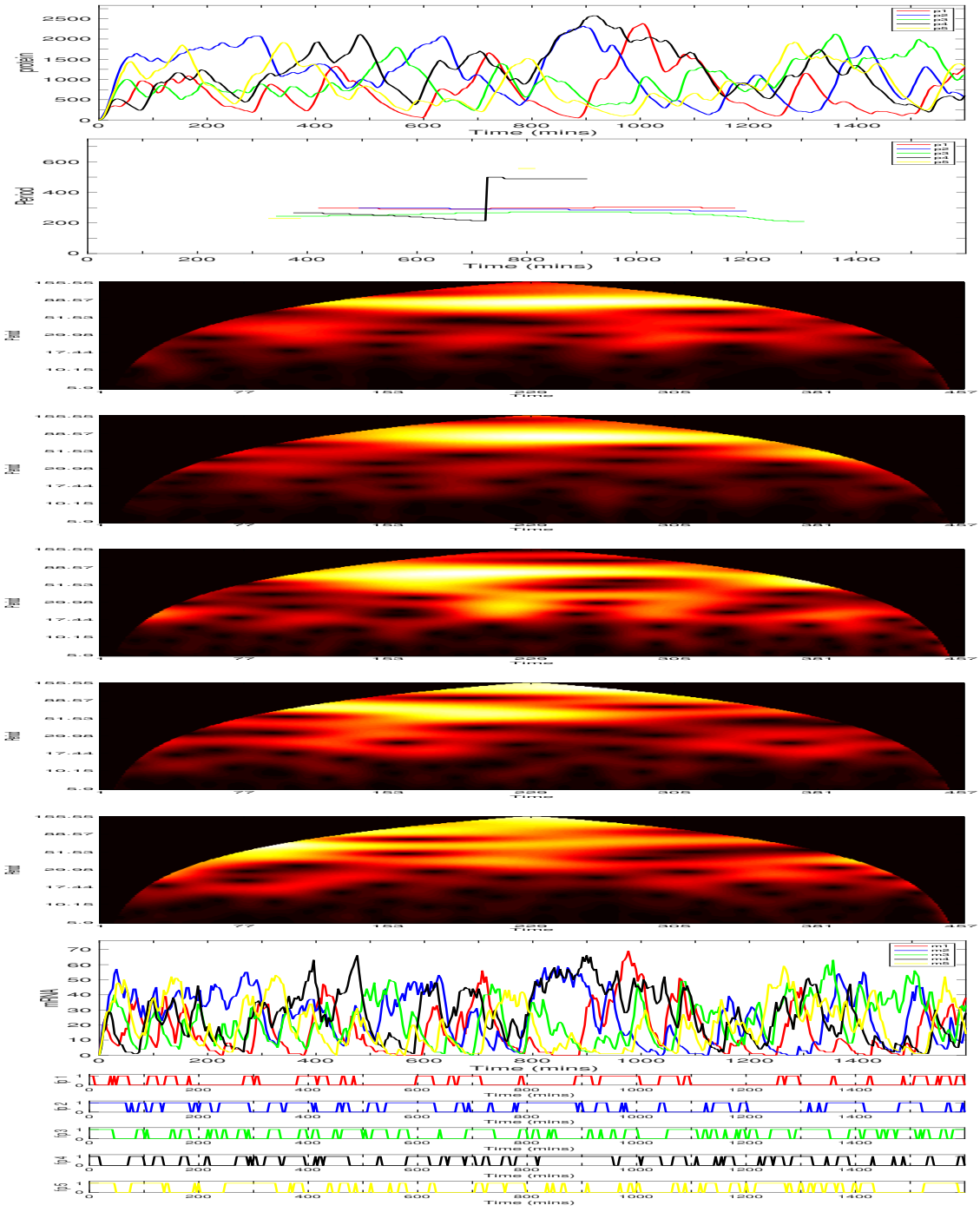

Figure 5.86: *The five gene repressilator with individual gene sites clustered together at the nuclear membrane. Each gene site lying on a vertex of a pentagon with sides* $1\mu m$ *long.* $g_1 = (0\mu m, gx + 2\mu m, 0\mu m)$, $g_2 = (gx\cos(\pi/10)\mu m, gx\sin(\pi/10) + 2\mu m, 0\mu m)$, $g_3 = (gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10) + 2\mu m, 0\mu m)$, $g_4 = (-gx\cos(3\pi/10)\mu m, -gx\sin(3\pi/10) + 2\mu m, 0\mu m)$ *and* $g_5 = (-gx\cos(\pi/10)\mu m, gx\sin(\pi/10)\mu m, 0\mu m)$, *where* $gx = 1/2\sin(\pi/5)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products and Gene5/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

For $D = 1 \times 10^{-12} m^2 min^{-1}$ the average period for 100 trajectories was $T = 386$, ranging between, $191 \leq T \leq 557$. The average protein number was $\bar{p} = 414$, ranging between $62 \leq \bar{p} \leq 829$. The average peak value of protein number was $\bar{p}_{peak} = 2473$, ranging between $495 \leq \bar{p}_{av} \leq 3901$.

For $D = 1 \times 10^{-10} m^2 min^{-1}$ the average period for 100 trajectories was $\bar{T} = 303$, ranging between, $158 \leq \bar{T} \leq 557$. The average protein number was $\bar{p} = 1894$, ranging between $1234 \leq \bar{p} \leq 2621$. The average peak value of protein number was $\bar{p}_{peak} = 4447$, ranging between $2904 \leq \bar{p}_{peak} \leq 5233$.

Once again, oscillatory behaviour seems elusive. We suggest that WAVOS only picks up the smaller modal periods of the system.

See 5.87 to see a summarising plot of protein copy number against the diffusion parameter. Generally, as diffusion is increased, the copy number of protein increases. In terms of gene site configurations, as the gene sites are moved closer to the nucleus membrane, the copy number increased, specifically for the slower diffusion regime. However, there is no significant difference in copy number due to gene site configuration for the faster diffusion regime.

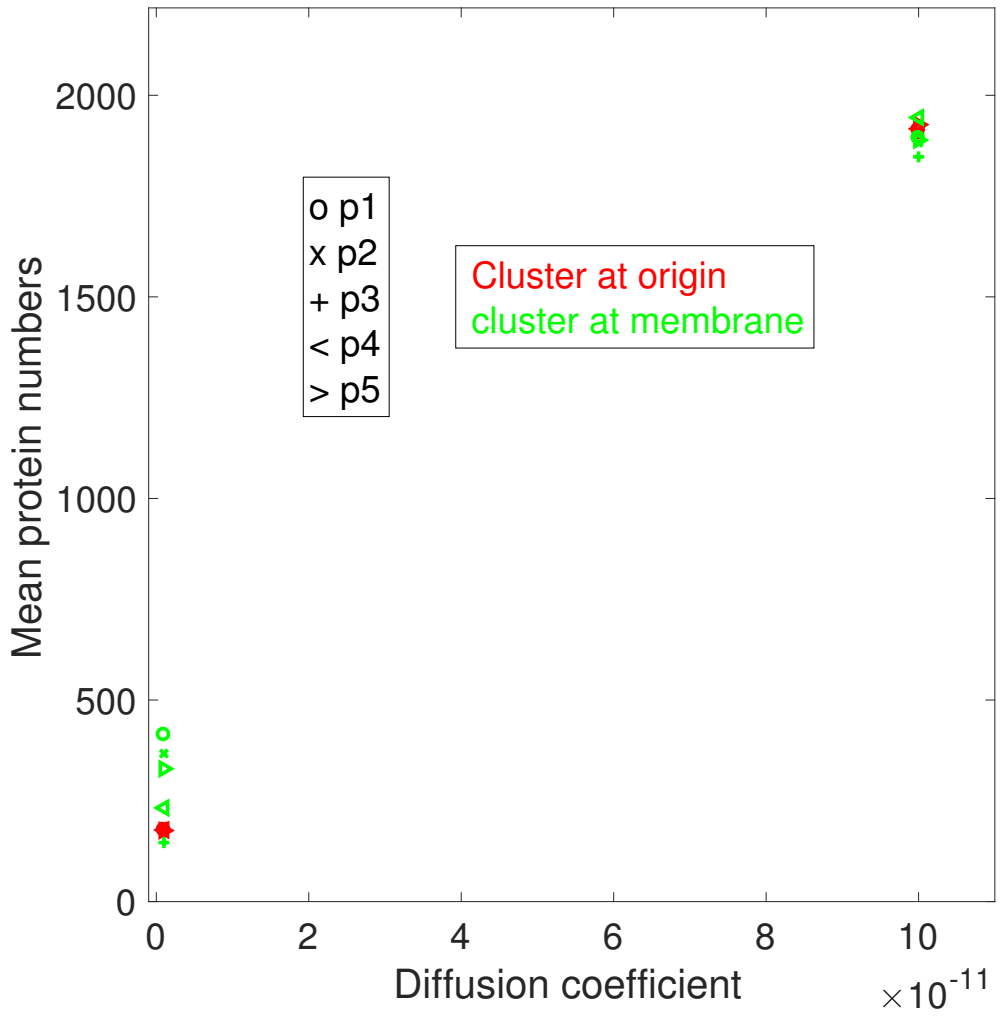

Figure 5.87: *Left: The mean value of protein calculated for the five-gene repressilator against the diffusion parameter. Right: The mean period calculated for the five-gene repressilator agains the diffusion parameter. Red denotes the gene site cluster at the origin, blue denotes the gene sites at the membrane on opposite ends of the x-axis and green denotes the cluster of genes close to the nuclear membrane.*

5.6 The six gene repressilator

Now we consider the six gene repressilator with individual gene sites clustered together at the origin. Each gene site lies on a vertex of a hexagon with coordinates, $g_1 = (0\mu m, 1\mu m, 0\mu m)$, $g_2 = (\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$, $g_3 = (\cos(\pi/3)\mu m, -\sin(\pi/3)\mu m, 0\mu m)$ $g_4 = (0\mu m, -1\mu m, 0\mu m)$, $g_5 = (-\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$ and $g_6 = (-\cos(\pi/3)\mu m, \sin(\pi/$

The first set of three simulations are for the smaller diffusion regime, $D = 1 \times 10^{-12}$ and the second set of three simulations are for the faster diffusion regime, $D = 1 \times 10^{-10}$.

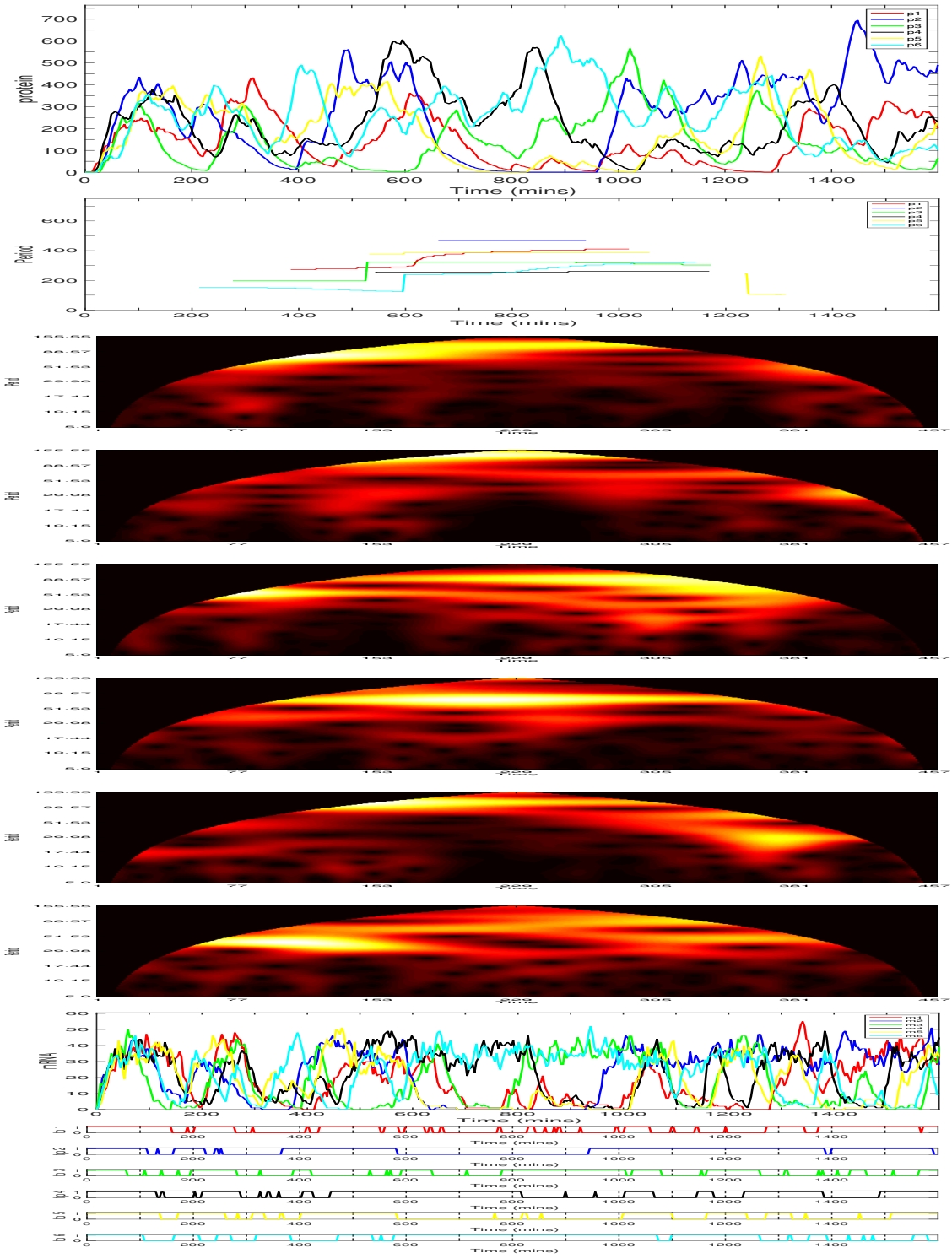

Figure 5.88: *The six gene repressilator with individual gene sites clustered together at the origin. Each gene site lying on a vertex of a hexagon with coordinates,* $g_1 = (0\mu m, 1\mu m, 0\mu m)$, $g_2 = (\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$, $g_3 = (\cos(\pi/3)\mu m, -\sin(\pi/3)\mu m, 0\mu m)$, $g_4 = (0\mu m, -1\mu m, 0\mu m)$, $g_5 = (-\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$ *and* $g_6 = (-\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products, Gene5/products and Gene6/products. Top to bottom: protein, period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

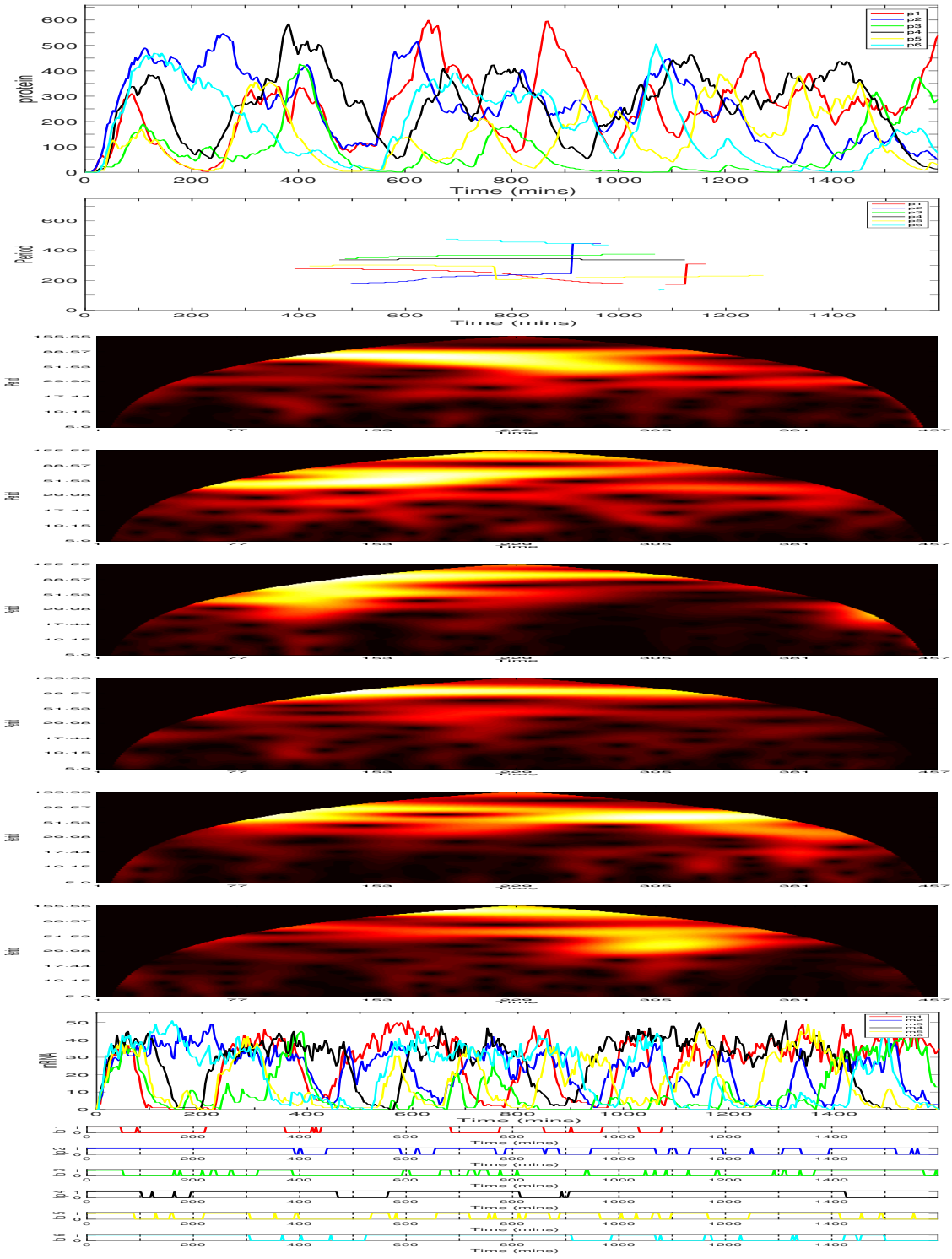

Figure 5.89: *The six gene repressilator with individual gene sites clustered together at the origin. Each gene site lying on a vertex of a hexagon with coordinates,* $g_1 = (0\mu m, 1\mu m, 0\mu m)$, $g_2 = (\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$, $g_3 = (\cos(\pi/3)\mu m, -\sin(\pi/3)\mu m, 0\mu m)$, $g_4 = (0\mu m, -1\mu m, 0\mu m)$, $g_5 = (-\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$ *and* $g_6 = (-\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products, Gene5/products and Gene6/products. Top to bottom: protein, period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

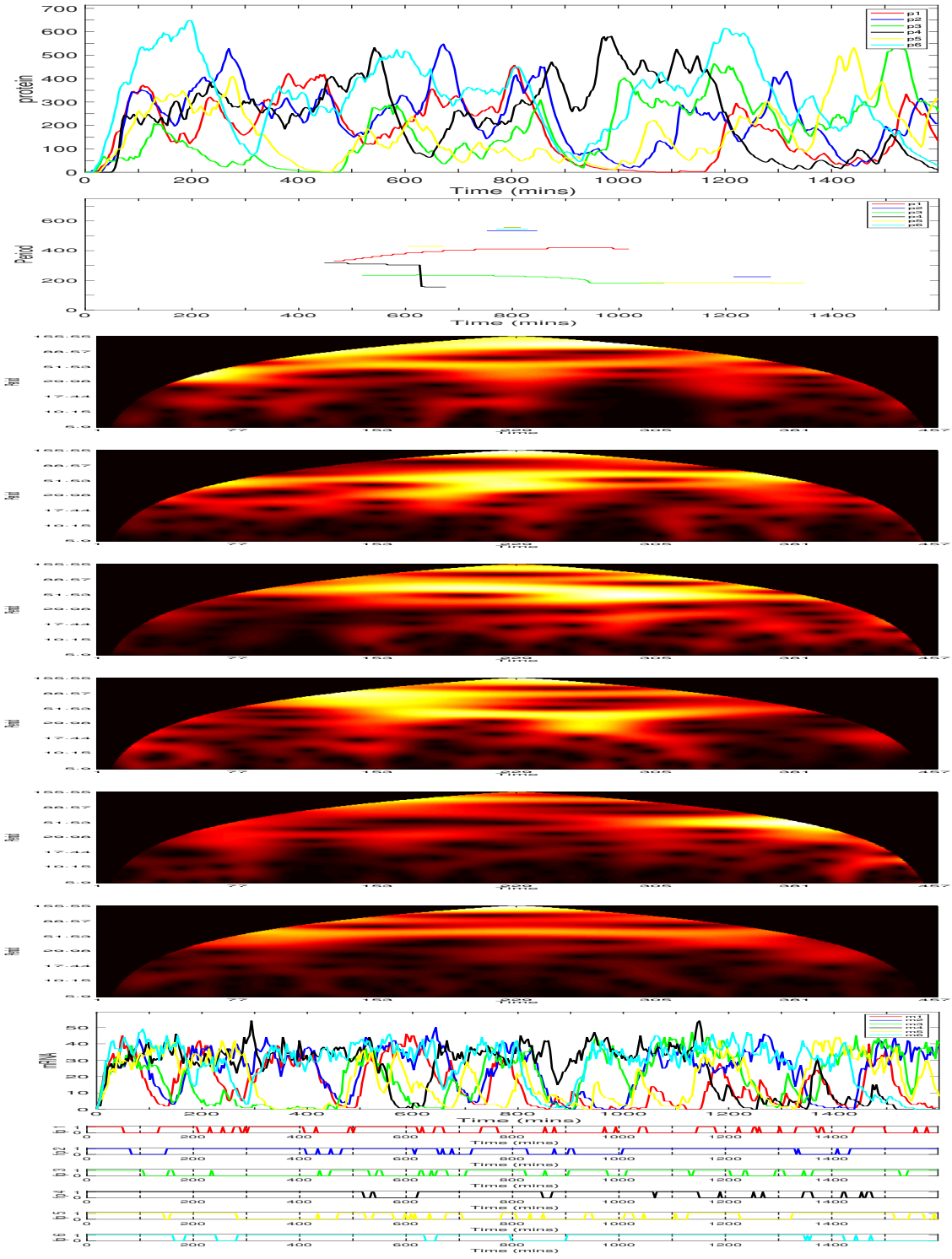

Figure 5.90: *The six gene repressilator with individual gene sites clustered together at the origin. Each gene site lying on a vertex of a hexagon with coordinates,* $g_1 = (0\mu m, 1\mu m, 0\mu m)$, $g_2 = (\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$, $g_3 = (\cos(\pi/3)\mu m, -\sin(\pi/3)\mu m, 0\mu m)$, $g_4 = (0\mu m, -1\mu m, 0\mu m)$, $g_5 = (-\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$ *and* $g_6 = (-\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products, Gene5/products and Gene6/products. Top to bottom: protein, period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

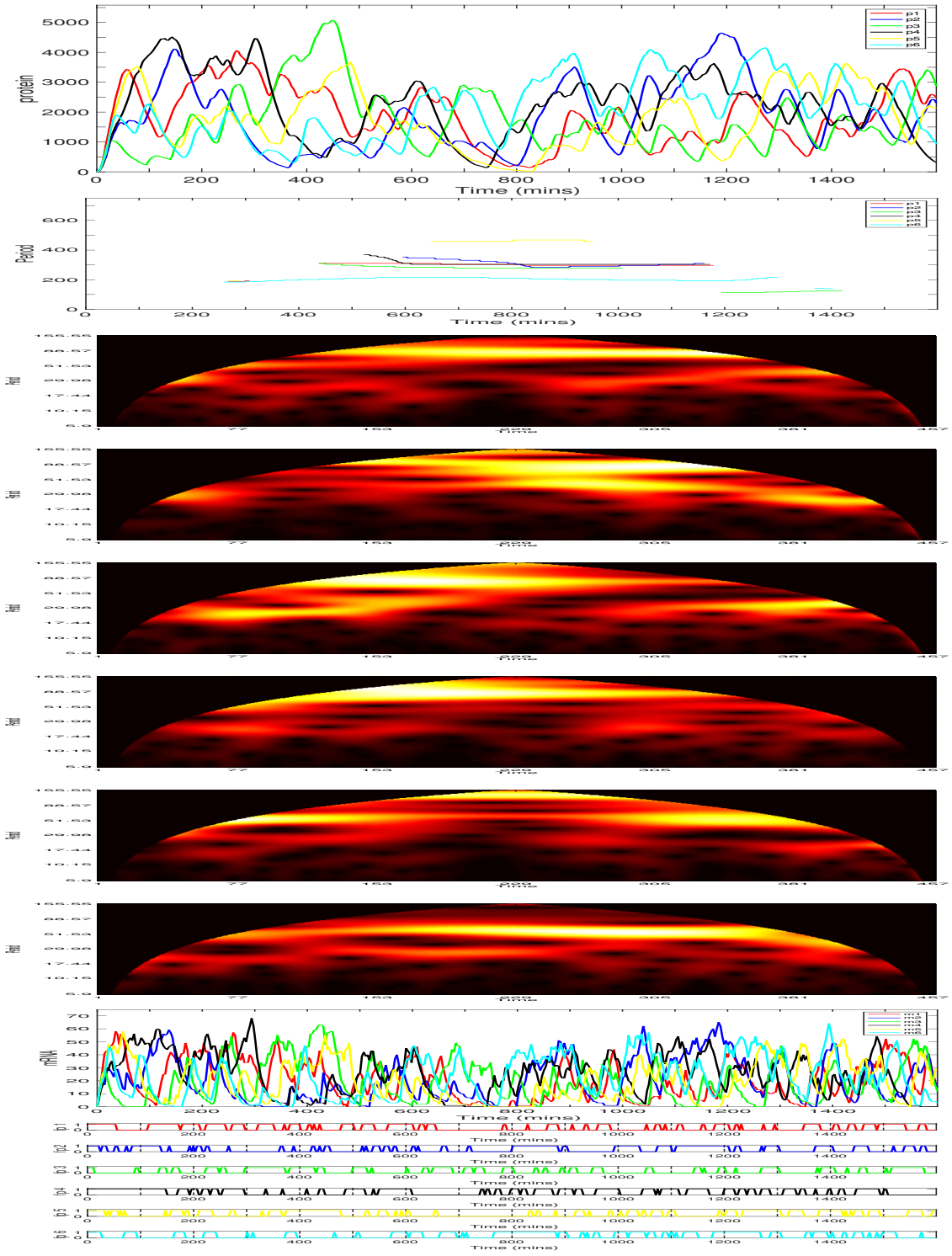

Figure 5.91: *The six gene repressilator with individual gene sites clustered together at the origin. Each gene site lying on a vertex of a hexagon with coordinates,* $g_1 = (0\mu m, 1\mu m, 0\mu m)$, $g_2 = (\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$, $g_3 = (\cos(\pi/3)\mu m, -\sin(\pi/3)\mu m, 0\mu m)$, $g_4 = (0\mu m, -1\mu m, 0\mu m)$, $g_5 = (-\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$ *and* $g_6 = (-\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products, Gene5/products and Gene6/products. Top to bottom: protein, period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

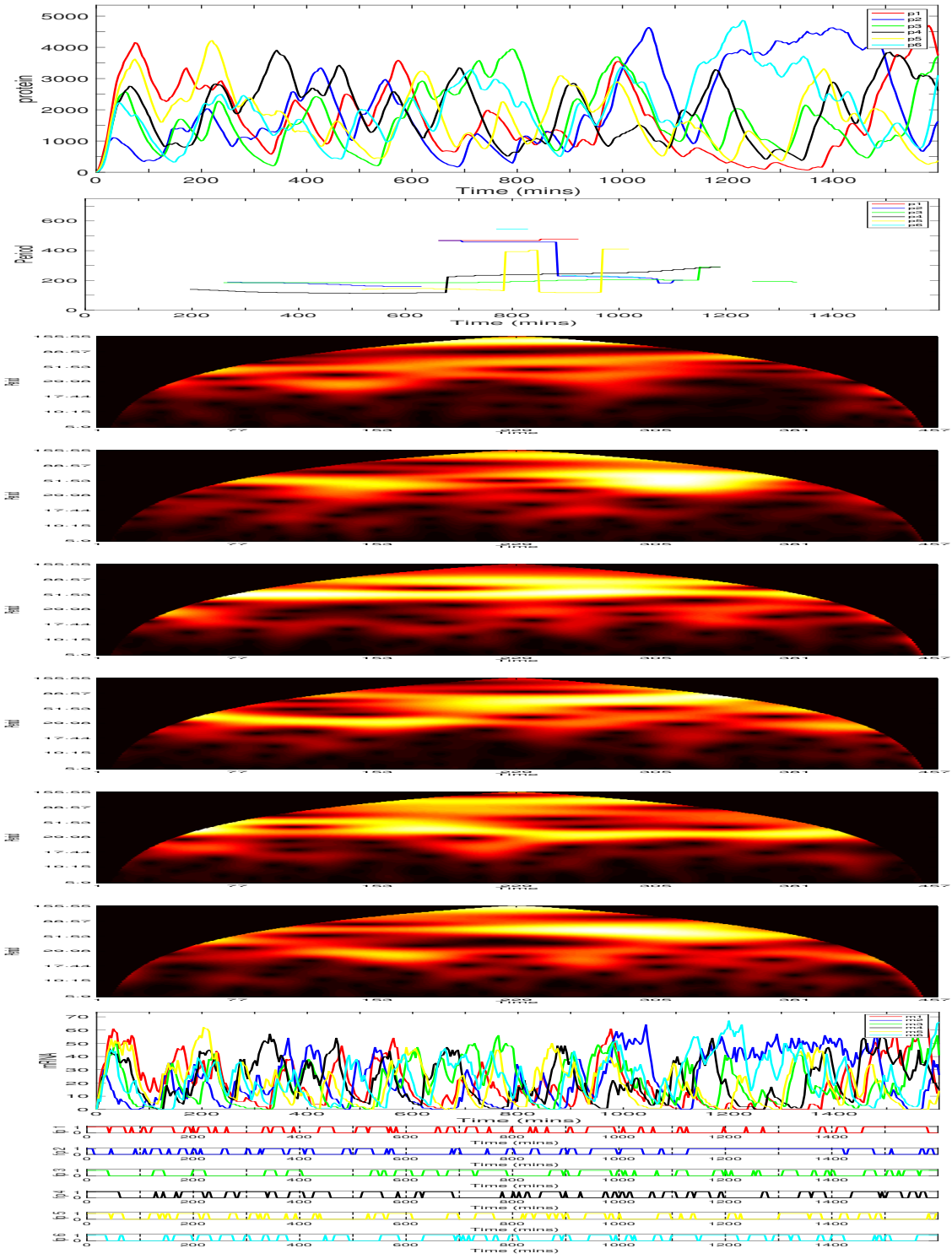

Figure 5.92: *The six gene repressilator with individual gene sites clustered together at the origin. Each gene site lying on a vertex of a hexagon with coordinates,* $g_1 = (0\mu m, 1\mu m, 0\mu m)$, $g_2 = (\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$, $g_3 = (\cos(\pi/3)\mu m, -\sin(\pi/3)\mu m, 0\mu m)$, $g_4 = (0\mu m, -1\mu m, 0\mu m)$, $g_5 = (-\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$ *and* $g_6 = (-\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products, Gene5/products and Gene6/products. Top to bottom: protein, period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

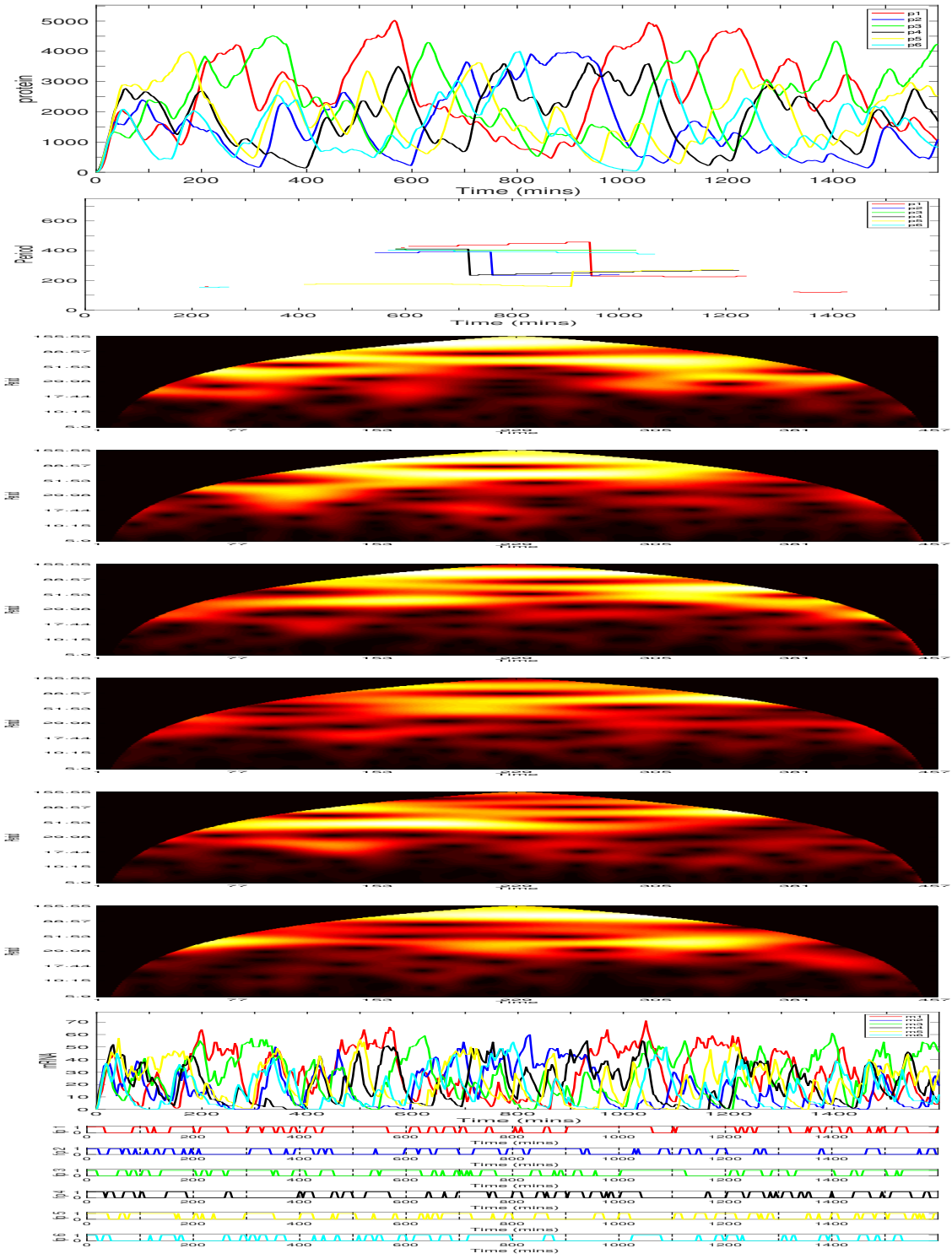

Figure 5.93: *The six gene repressilator with individual gene sites clustered together at the origin. Each gene site lying on a vertex of a hexagon with coordinates,* $g_1 = (0\mu m, 1\mu m, 0\mu m)$, $g_2 = (\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$, $g_3 = (\cos(\pi/3)\mu m, -\sin(\pi/3)\mu m, 0\mu m)$, $g_4 = (0\mu m, -1\mu m, 0\mu m)$, $g_5 = (-\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$ *and* $g_6 = (-\cos(\pi/3)\mu m, \sin(\pi/3)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products, Gene5/products and Gene6/products. Top to bottom: protein, period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

For $D = 1 \times 10^{-12} m^2 min^{-1}$ the average period for 100 trajectories was $\bar{T} = 365$, ranging between, $215 \leq \bar{T} \leq 557$. The average protein number was $\bar{p} = 1937$, ranging between $100 \leq \bar{p} \leq 310$. The average peak value of protein number was $\bar{p}_{peak} = 536$, ranging between $381 \leq \bar{p}_{peak} \leq 779$.

For $D = 1 \times 10^{-10} m^2 min^{-1}$ the average period for 100 trajectories was $\bar{T} = 313$, ranging between, $123 \leq \bar{T} \leq 557$. The average protein number was $\bar{p} = 197.3$, ranging between $107.5 \leq \bar{p} \leq 272.3$. The average peak value of protein number was $\bar{p}_{peak} = 1958$, ranging between $1012 \leq \bar{p}_{av} \leq 2631$.

Now we consider the six gene repressilator with individual gene sites clustered together at the nuclear membrane. Each gene site is lying on a vertex of a hexagon with coordinates, $g_1 = (x_1, y_1, z_1) = (0\mu m, 2.85\mu m, 0\mu m)$, $(x_2, y_2, z_2) = g_2 = (\cos(\pi/6)\mu m, 1.85 + \sin(\pi/6)\mu m, 0\mu m)$, $g_3 = (x_3, y_3, z_3) = (\cos(\pi/6)\mu m, 1.85 - \sin(\pi/6)\mu m, 0\mu m)$, $g_4 = (x_4, y_4, z_4) = (0\mu m, 0.85\mu m, 0\mu m)$, $g_5 = (x_5, y_5, z_5) = (-\cos(\pi/6)\mu m, 1.85 - \sin(\pi/6)\mu m, 0\mu m)$ and $g_6 = (x_6, y_6, z_6) = (-\cos(\pi/6)\mu m, 1.85 + \sin(\pi/6)\mu m, 0\mu m)$

The first set of three simulations are for the smaller diffusion regime, $D = 1 \times 10^{-12}$ and the second set of three simulations are for the faster diffusion regime, $D = 1 \times 10^{-10}$.

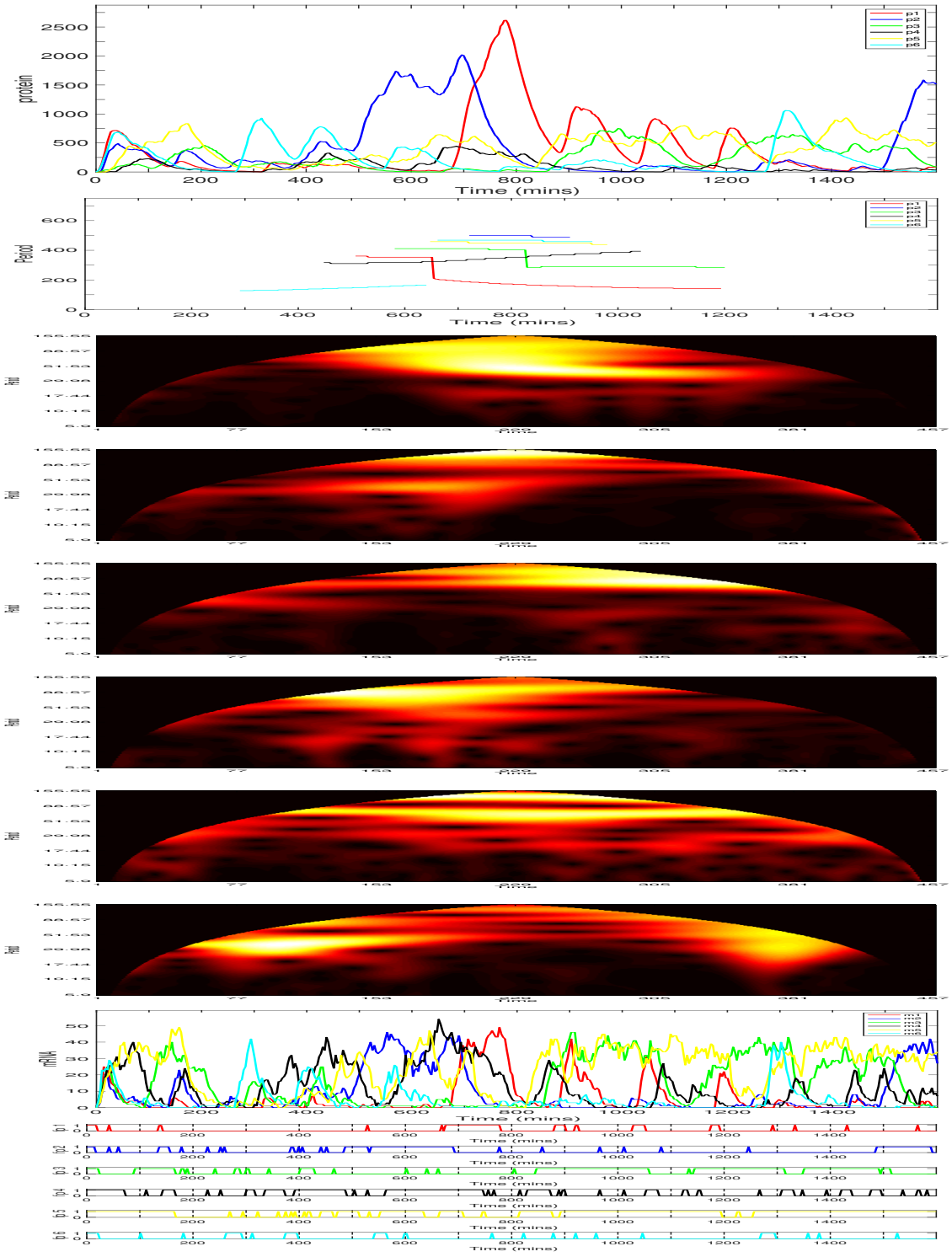

Figure 5.94: *The six gene repressilator with individual gene sites clustered together at the nucleus membrane. Each gene site lying on a vertex of a hexagon with coordinates,* $g_1 = (0\mu m, 2.85\mu m, 0\mu m)$, $g_2 = (\cos(\pi/3)\mu m, 1.85 + \sin(\pi/3)\mu m, 0\mu m)$, $g_3 = (\cos(\pi/3)\mu m, 1.85 - \sin(\pi/3)\mu m, 0\mu m)$, $g_4 = (0\mu m, 0.85\mu m, 0\mu m)$, $g_5 = (-\cos(\pi/3)\mu m, 1.85 - \sin(\pi/3)\mu m, 0\mu m)$ *and* $g_6 = (-\cos(\pi/3)\mu m, 1.85 + \sin(\pi/3)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products, Gene5/products and Gene6/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

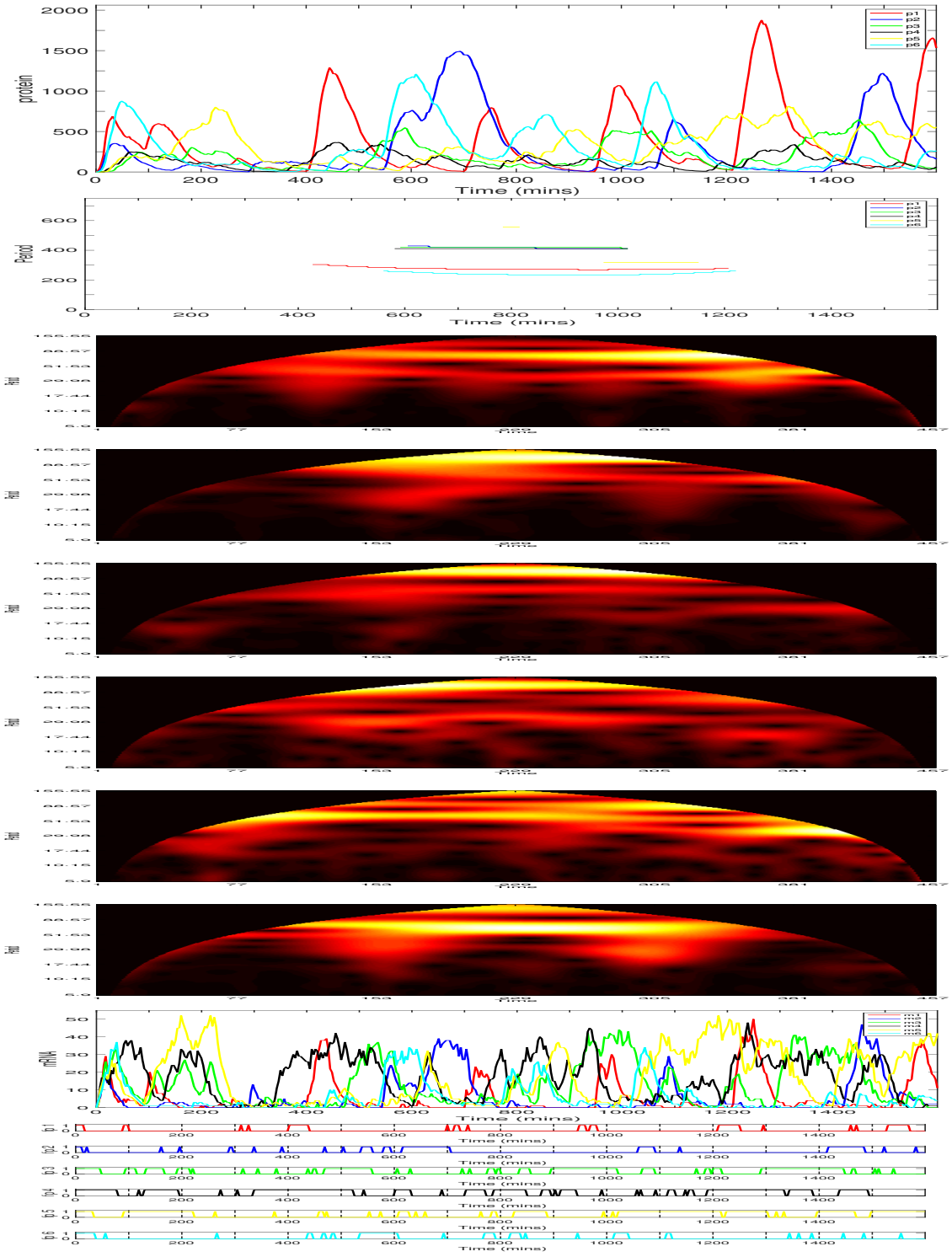

Figure 5.95: *The six gene repressilator with individual gene sites clustered together at the nucleus membrane. Each gene site lying on a vertex of a hexagon with coordinates,* $g_1 = (0\mu m, 2.85\mu m, 0\mu m)$, $g_2 = (\cos(\pi/3)\mu m, 1.85 + \sin(\pi/3)\mu m, 0\mu m)$, $g_3 = (\cos(\pi/3)\mu m, 1.85 - \sin(\pi/3)\mu m, 0\mu m)$, $g_4 = (0\mu m, 0.85\mu m, 0\mu m)$, $g_5 = (-\cos(\pi/3)\mu m, 1.85 - \sin(\pi/3)\mu m, 0\mu m)$ *and* $g_6 = (-\cos(\pi/3)\mu m, 1.85 + \sin(\pi/3)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products, Gene5/products and Gene6/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

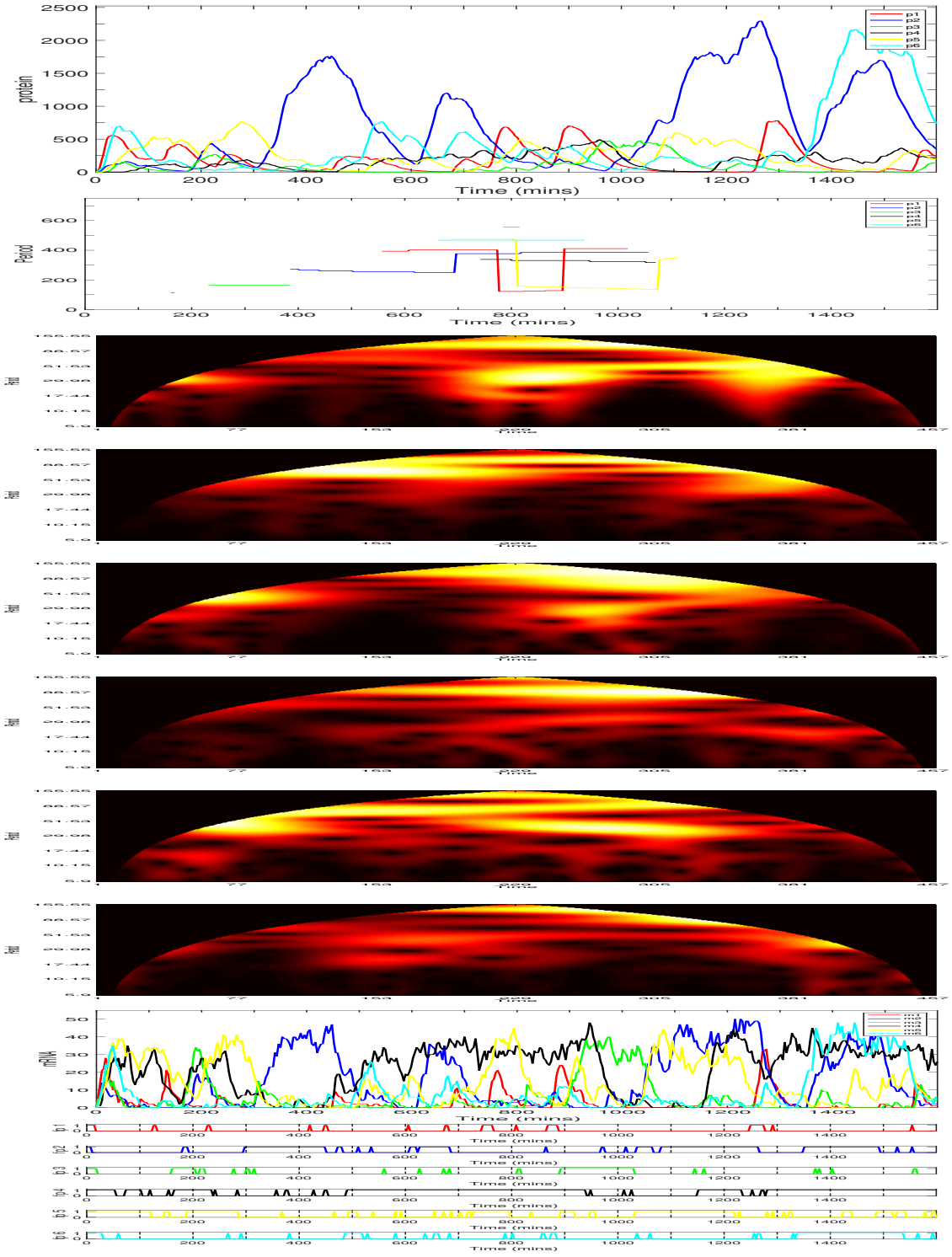

Figure 5.96: *The six gene repressilator with individual gene sites clustered together at the nucleus membrane. Each gene site lying on a vertex of a hexagon with coordinates,* $g_1 = (0\mu m, 2.85\mu m, 0\mu m)$, $g_2 = (\cos(\pi/3)\mu m, 1.85 + \sin(\pi/3)\mu m, 0\mu m)$, $g_3 = (\cos(\pi/3)\mu m, 1.85 - \sin(\pi/3)\mu m, 0\mu m)$, $g_4 = (0\mu m, 0.85\mu m, 0\mu m)$, $g_5 = (-\cos(\pi/3)\mu m, 1.85 - \sin(\pi/3)\mu m, 0\mu m)$ *and* $g_6 = (-\cos(\pi/3)\mu m, 1.85 + \sin(\pi/3)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-12} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products, Gene5/products and Gene6/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

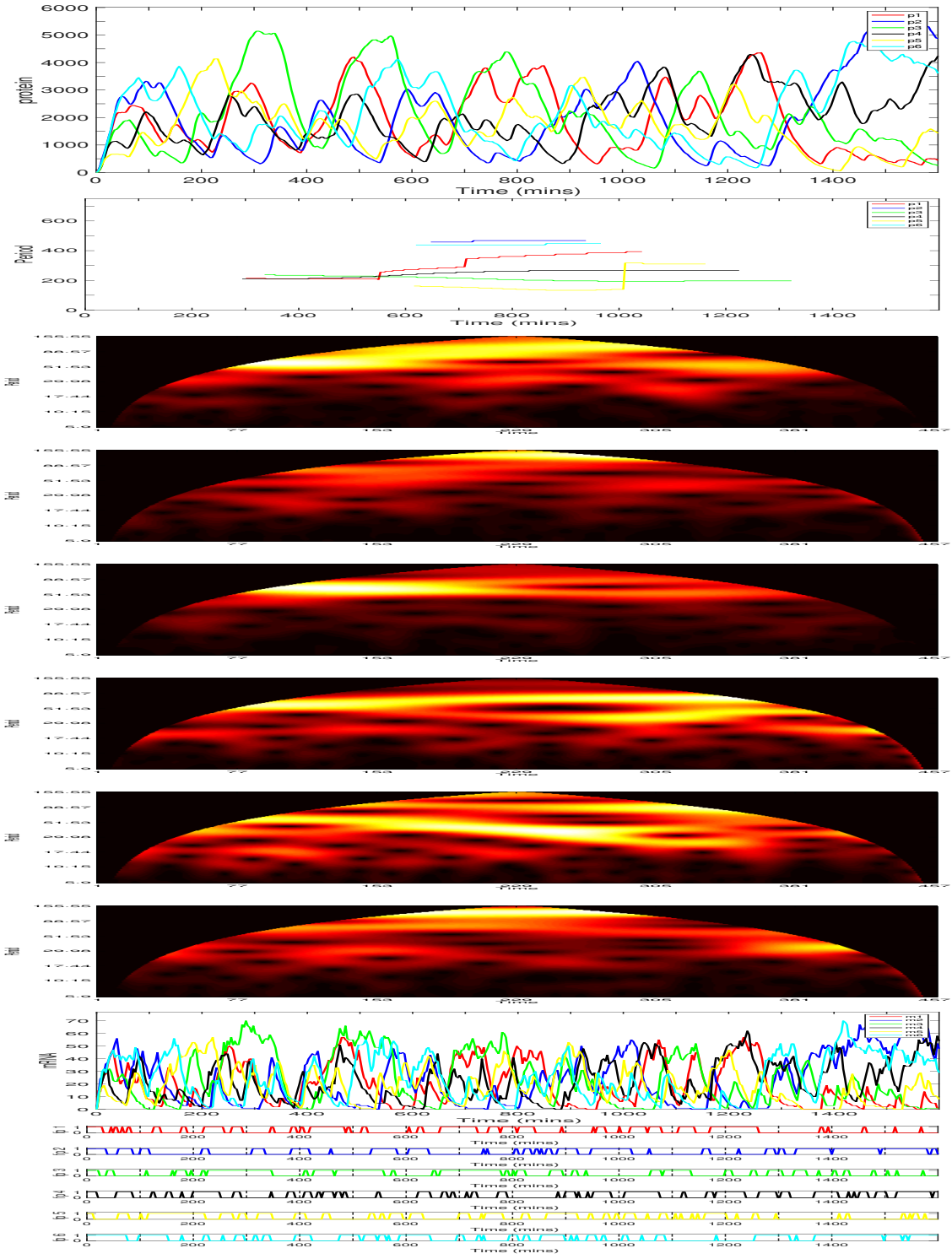

Figure 5.97: *The six gene repressilator with individual gene sites clustered together at the nucleus membrane. Each gene site lying on a vertex of a hexagon with coordinates,* $g_1 = (0\mu m, 2.85\mu m, 0\mu m)$, $g_2 = (\cos(\pi/3)\mu m, 1.85 + \sin(\pi/3)\mu m, 0\mu m)$, $g_3 = (\cos(\pi/3)\mu m, 1.85 - \sin(\pi/3)\mu m, 0\mu m)$, $g_4 = (0\mu m, 0.85\mu m, 0\mu m)$, $g_5 = (-\cos(\pi/3)\mu m, 1.85 - \sin(\pi/3)\mu m, 0\mu m)$ *and* $g_6 = (-\cos(\pi/3)\mu m, 1.85 + \sin(\pi/3)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products, Gene5/products and Gene6/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

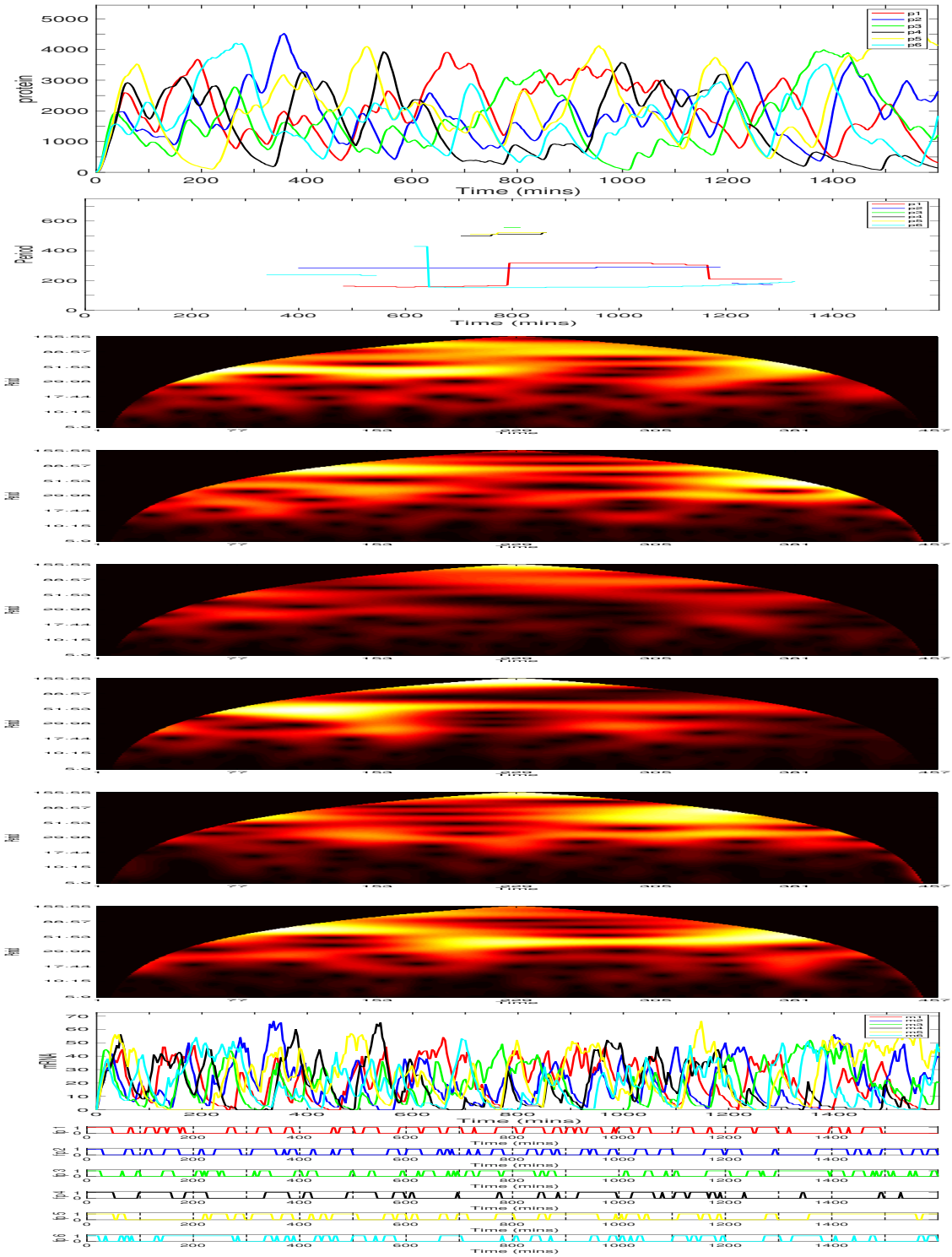

Figure 5.98: *The six gene repressilator with individual gene sites clustered together at the nucleus membrane. Each gene site lying on a vertex of a hexagon with coordinates,* $g_1 = (0\mu m, 2.85\mu m, 0\mu m)$, $g_2 = (\cos(\pi/3)\mu m, 1.85 + \sin(\pi/3)\mu m, 0\mu m)$, $g_3 = (\cos(\pi/3)\mu m, 1.85 - \sin(\pi/3)\mu m, 0\mu m)$, $g_4 = (0\mu m, 0.85\mu m, 0\mu m)$, $g_5 = (-\cos(\pi/3)\mu m, 1.85 - \sin(\pi/3)\mu m, 0\mu m)$ *and* $g_6 = (-\cos(\pi/3)\mu m, 1.85 + \sin(\pi/3)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products, Gene5/products and Gene6/products. From top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

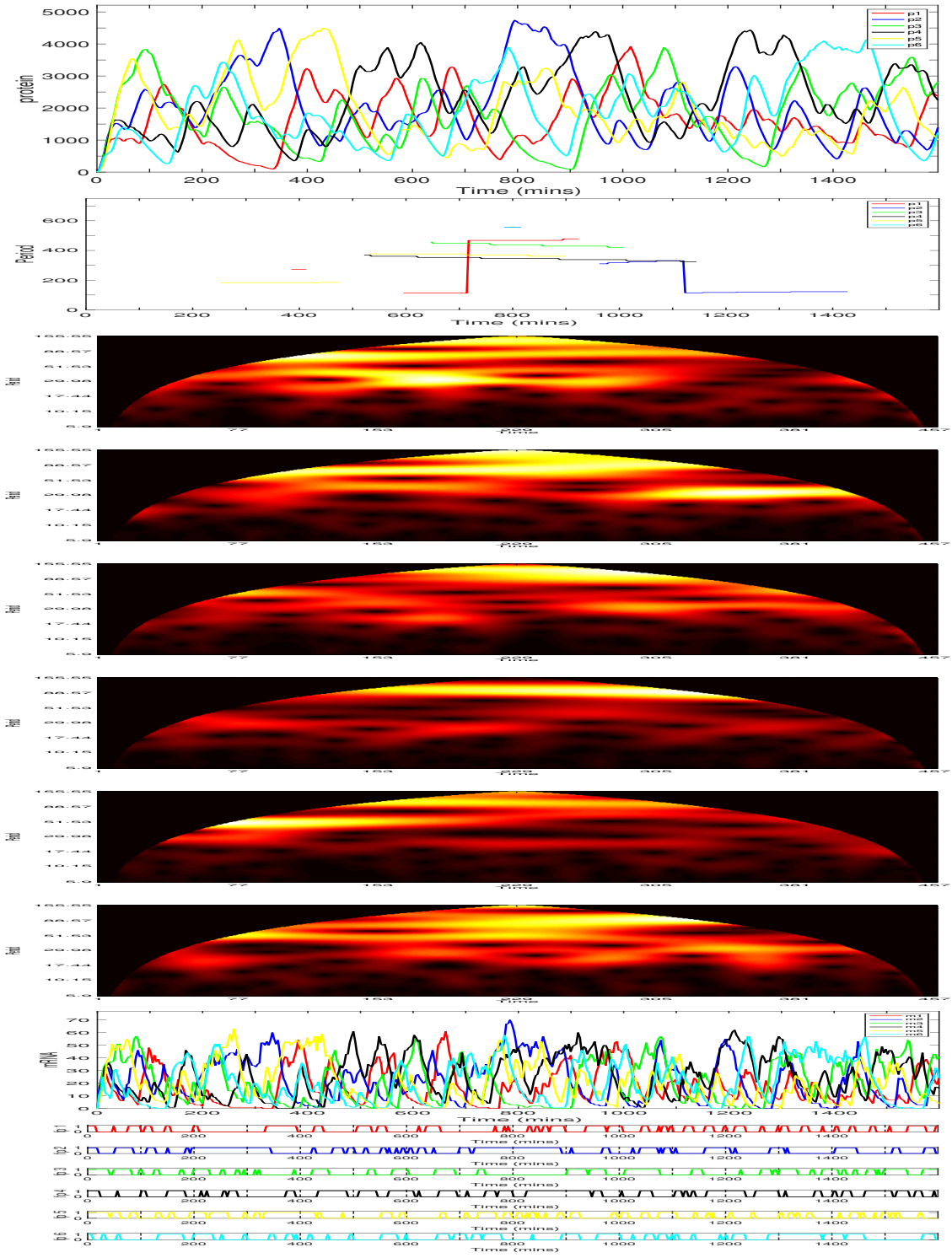

Figure 5.99: *The six gene repressilator with individual gene sites clustered together at the nucleus membrane. Each gene site lying on a vertex of a hexagon with coordinates,* $g_1 = (0\mu m, 2.85\mu m, 0\mu m)$, $g_2 = (\cos(\pi/3)\mu m, 1.85 + \sin(\pi/3)\mu m, 0\mu m)$, $g_3 = (\cos(\pi/3)\mu m, 1.85 - \sin(\pi/3)\mu m, 0\mu m)$, $g_4 = (0\mu m, 0.85\mu m, 0\mu m)$, $g_5 = (-\cos(\pi/3)\mu m, 1.85 - \sin(\pi/3)\mu m, 0\mu m)$ *and* $g_6 = (-\cos(\pi/3)\mu m, 1.85 + \sin(\pi/3)\mu m, 0\mu m)$. *A diffusion coefficient* $D = 1 \times 10^{-10} m^2 min^{-1}$. *Species are colour coded dependent on genetic relation: Gene1/products, Gene2/products, Gene3/products, Gene4/products, Gene5/products and Gene6/products. Top to bottom: Protein, Period, mRNA and free promoter time series. Each time unit (along the x-axis) of each heatmap is equal to 3.5 minutes. This scaling of the x-axis in turn affects the scaling of the period (y-axis, of the heatmaps, which has a maximum value of approximately one third of the given time span) and thus the y-axis of each heatmap is down scaled by a factor of 3.5. Translation allowed to occur equally over the entire cytoplasm. Initial conditions are each promoter is free.*

For $D = 1 \times 10^{-12} m^2 min^{-1}$ the average period for 100 trajectories was $\bar{T} = 366$, ranging between, $203 \leq \bar{T} \leq 557$. The average protein number was $\bar{p} = 463$, ranging between $151 \leq \bar{p} \leq 886$. The average peak value of protein number was $\bar{p}_{peak} = 2348$, ranging between $781 \leq \bar{p}_{peak} \leq 3388$.

For $D = 1 \times 10^{-10} m^2 min^{-1}$ the average period for 100 trajectories was $\bar{T} = 303$, ranging between, $180 \leq \bar{T} \leq 557$. The average protein number was $\bar{p} = 1904$, ranging between $1035 \leq \bar{p} \leq 2833$. The average peak value of protein number was $\bar{p}_{peak} = 4369$, ranging between $2220 \leq \bar{p}_{av} \leq 5803$.

Once again, oscillatory behaviour seems elusive. We suggest that WAVOS only picks up the smaller modal periods of the system.

See 5.100 to see a summarising plot of protein copy number against the diffusion parameter. Generally, as diffusion is increased, the copy number of protein increases.

5.7 Discussion

As noted in the previous chapter, in the biological world, state transitions (i.e. the transformation of a biological system from one state to another) are vitally important in many key biological systems, including intracellular signalling pathways. For example, the cell cycle with its four distinct phases: G1, S, G2 and M phases; cell differentiation contains different stages like cell proliferation, growth arrest and mature differentiation; and cancer development mainly involves three steps as mutation, promotion and invasion. A fundamental but important question is how to trace the temporal characteristics or dynamics of a biological system during a particular state transition process (Zeng & Chen 2012).

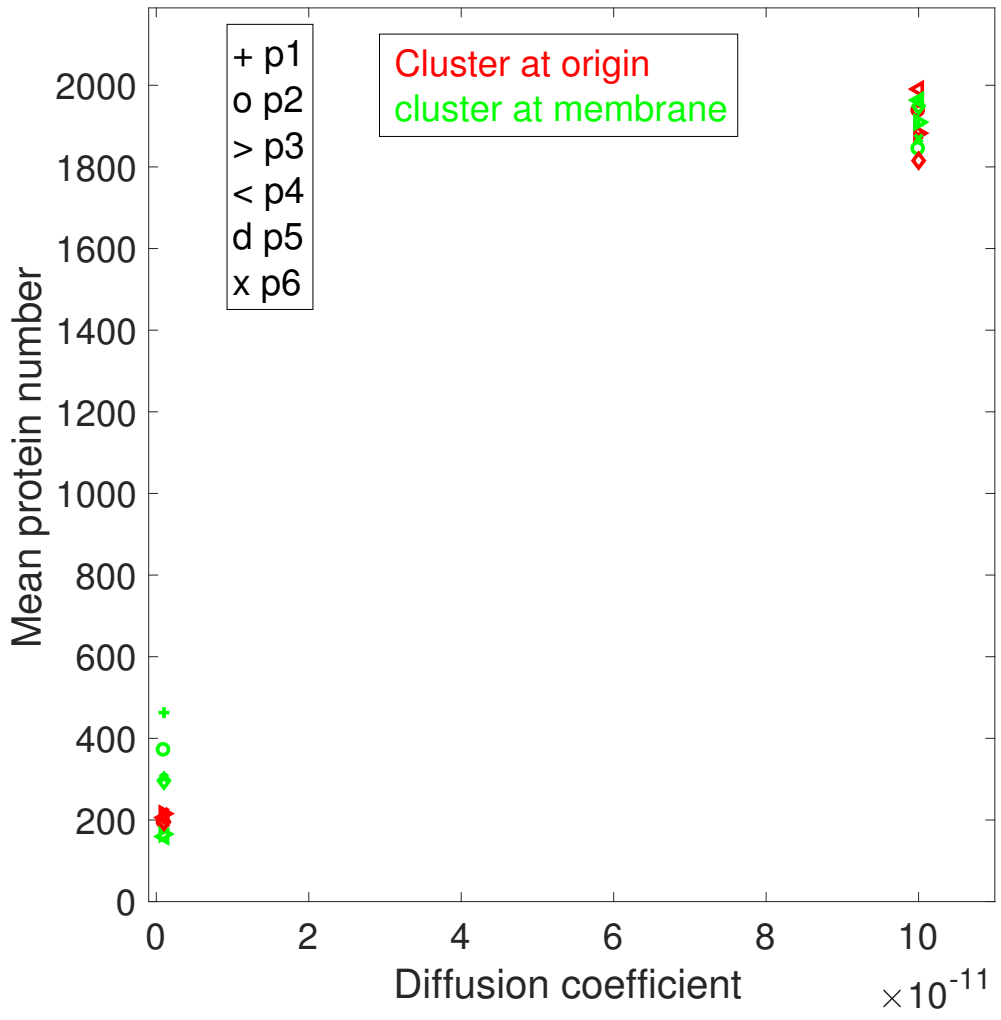

Figure 5.100: *Left: The mean value of protein calculated for the six-gene repressilator against the diffusion parameter. Right: The mean period calculated for the six-gene repressilator agains the diffusion parameter. Red denotes the gene site cluster at the origin, blue denotes the gene sites at the membrane on opposite ends of the x-axis and green denotes the cluster of genes close to the nuclear membrane.*

In our spatial stochastic model of the synthetic n-gene repressilator, we saw multi-modal distributions of periods of oscillation. In cellular processes, different mechanisms are controlled on different time scales. We suggest that there is a greater complexity involved in GRNs. That the oscillatory might act as clocks on different time scales. The low modal periods may act as transients in the system, while the larger modes, may control a mechanism on a longer time scale.

Changes in spatial parameters, such as the diffusion coefficient and the location of gene sites definitely had a great effect on the dynamics, specifically in terms of molecular numbers. We stress the vital importance using spatial models in the simulation of GRNs.

Chapter 6

The NFκB and HIF1α Gene Regulatory Network

6.1 Introduction

Nuclear factor κB (NFκB) was discovered in 1986 as a nuclear factor in B lymphocytes, regulating the gene encoding the immunoglobulin κ light polypeptide chain, hence its given name. NFκB has since been found to be present in almost all mammalian cell types and is activated in response to many different stimuli, including environmental cues such as hypoxia and ultraviolet radiation; infectious agents, such as bacteria and virus, and extra- and intracellular stress, such as inflammatory cytokines and DNA damage. Due to the diversity of the means to NFκB activation, it is not surprising that NFκB has been found to have the potential to control the transcriptional activity of over three hundred genes and to thus play a role in many different processes. For example, NFκB has been found to be a major player in coordinating processes such as adaptive and innate immunity, development and cell survival. In one respect,

NFκB is a form of guardian that is normally held latent in the cytoplasm due to its sequestration by the inhibitor of NFκB (IκB) family, to which it is readily released and transiently activated at the first instance the cell senses certain environmental changes that are a threat to cellular homeostasis and survival. This transient mechanism is enabled as the genes encoding IκB and other negative regulators of NFκB are early transcriptional targets of NFκB.

However, due to the role NFκB plays in processes such as immunity and survival it is not surprising that its dysregulation in mammalian cells might have hazardous implications. When phenotypes, resulting largely due to NFκB activation, such as the inflammatory response, angiogenesis, apoptosis and cell proliferation become chronic and out of control, the onset of diseases such as chronic inflammatory diseases, autoimmune diseases and the initiation and progression of cancer are almost inevitable. The NFκB pathway can also, unfortunately, be activated during chemotherapy to counteract the intended chemotherapeutic effects and once again aid in cell survival and hence tumour progression. Although, it has recently become apparent that the effects of cytotoxic drugs on NFκB function are also quite diverse and may differ between cell and tumor types or even the different stages of tumor development.

6.2 The NFκB family

NFκB is a dimeric transcription factor composed as a homo- or heterodimer from the following family of subunits: RelA (p65), RelB, cRel, NFκB1 (p50) and NFκB2 (p52), the latter two subunits being proteolytically processed from p105 and p100 respectively. Each NFκB subunit is related via their commonality of the Rel Homology Domain (RHD) in their N termini. The RHD contains sequences required for nuclear

localisation; NFκB subunit dimerisation; DNA binding and association with the inhibitor IκB family. The members of the IκB family are IκBα, IκBβ, IκBε, IκBζ and Bcl3 (B cell lymphoma 3) and each share the structural feature of containing 5-7 ankyrin repeats, which when bound to an NFκB dimer masks NFκB's nuclear localisation sequence (NLS). Each IκB member has different binding affinities to different NFκB subunit members, contributing further to the diversity of NFκB activation and its functioning. The NFκB family can be categorised further into two subfamilies, one containing RelA, RelB and cRel and the other NFκB1 and NFκB2, (p105 and p100). The first subfamily contain a transcriptional activation domain (TAD), the latter subfamily do not. The subunits NFκB1 and NFκB2 can only contribute to transcriptional activation if they are in an NFκB dimer with a member of the TAD subfamily but if not they can act as transcriptional repressors. Also, like the IκB family, p100 and p105 contain ankyrin repeats, which classifies them further as overlapping members of the IκB family, as their binding to an NFκB subunit with a TAD shields the NLS, resulting in the NFκB dimer being kept held inactive in the cytoplasm. The processing of p105 to NFκB1 mainly occurs constitutively but the processing of p100 to NFκB2 mainly only happens due to the presence of specific extracellular ligands binding to specific cell receptors that trigger the activation of the NFκB inducing kinase (NIK) that triggers this process. Following the induced proteolytic processing of p100, RelB/NFκB2 dimerisation and nuclear translocation occur. This cascade is known as the non canonical (alternative) NFκB signalling pathway. There are two other major pathways that lead to NFκB activation: the atypical and canonical (classical) pathways. The atyptical pathways are activated independently of external ligand binding and instead are mainly activated due to signalling occuring via internal cellular stresses. The canonical pathway is the pathway most commonly seen to be activated within most cell types and unlike the non canonical pathway is activated due to external ligand/cell receptor binding. Once an NFκB pathway is activated, NFκB travels into the nucleus, where it

can then bind to its target genes. There are many factors that are believed to contribute to NFκB gene selectivity and transcriptional control.

6.2.1 Proposed mechanisms of gene selection

Taking all mammalian cell types into account, NFκB has been found to have the potential to transcriptionally control over 300 genes, (see www.nf-kb.org).

Due to NFκB's ability to influence expression of hundreds of genes (and the fact that its dysregulation can lead to diseases), the activity of NFκB is tightly regulated on multiple levels, such that the desired outcome, in response to a signal is achieved. Cells must decipher multiple signals, along with stochastic noise and encode this information in its genetic networks to achieve the required phenotypic response. The required phenotypic response is produced by the activation or repression of a specific subset of genes.

The question then arises, what is the mechanism behind the selection of NFκB dependent gene activation.

Which specific subset of genes NFκB has been found to control is both stimulus and cell type dependent.

It is cell type dependent due to biological structure formed throughout cellular development, such as the deposition of chromatin barriers.

Intimately linked to cell type dependence, is stimulus dependence because of the epigenetic inducing information the stimulus may deliver to the cell which may enable the overcoming of certain chromatin barriers.

Not only is the response dependent on the specific stimulus type but also on the temporal profile of that given stimulus.

Furthermore, the specific genetic response is also dependent on the dynamics of NFκB nuclear-cytoplasmic oscillations that are produced due to negative feedback from early response genes (Ashall et al. 2009, Lee et al. 2014, Nelson et al. 2004). The nuclear-cytoplasmic oscillation observed in experiments may enclose information in the form of a genetic clock. The nuclear-cytoplasmic oscillations may have functional consequences via the transfer of analogue information in the form of response time, oscillatory period, amplitude and oscillation number but also on how oscillation persistence, may involve cycles of phosphorylation and dephosphorylation or other forms of post-translational modifications.

Furthermore, the posttranslational modification sites of NFκB are important for gene activation and crosstalk with other signaling pathways, which may also affect specificity. NFκB cooperates with multiple other signaling molecules and pathways. Prominent nodes of crosstalk are mediated by other transcription factors such as HIF (Hypoxia Inducible Factor), Notch, IRF and p53. These transcription factors either directly interact with NFκB subunits, deactivate NF$-\kappa$B activators or affect NFκB target genes.

Another regulatory layer may be through the combinatorial associations of the NFκB protein family members and how each dimer has differential preferences for variations of the DNA-binding sequence. Thus different target genes are differentially induced by distinct NFκB dimers.

In theory all dimers are predicted to exist but only 15 different dimers have been recorded so far and the physiological existence and relevance for all possible dimeric complexes has not yet been fully demonstrated. However, most dimers are rarely observed. The predominant NFκB dimers in most cell types is the RelA-NFκB1 heterodimer (Nelson et al. 2004, O'Dea & Hoffmann 2010).

Thus, together all these factors may contribute to the output signal that determines

specific genetic activity and hence, phenotypic response.

An increased understanding of the different roles and regulations of NFκB will create opportunities for the development of techniques that will allow existing and new therapeutic interventions for diseases, such as cancer to be more effective.

Thus, better predictions of the effect of a stimulus on the NFκB signalling pathway and overlapping crosstalk via mathematical modelling is essential.

6.3 The canonical NFκB signalling pathway

In the canonical NFκB signalling pathway, NFκB,

(predominantly experimentally observed to be RelA-NFκB1 heterodimer) is held inactive in the cytoplasm of unstimulated cells by the family of inhibitor IκB proteins, (predominantly IκBα). The binding of IκB to NFκB masks NFκB's NLS, which in turn prevents the binding of NFκB to nuclear pore complexes and hence nuclear translocation (O'Dea & Hoffmann 2010). IκB further spatially regulates NFκB by actively translocating to the nucleus, binding to nuclear NFκB and transporting it back to the cytoplasm (Arenzana-Seisdedos et al. 1997).

Particular examples of extracellular stimuli that activate the canonical NFκB signalling pathway are the proinflammatory cytokine: tumor necrosis factor α (TNFα) and the bacterial product: lipopolysaccharide (LPS). Upon ligand binding to a specific cellular membrane receptor, such as tumor necrosis factor α receptor 1 (TNFR1) or Toll like receptor 4 (TLR4), adaptor molecules, kinases and ubiquitin ligases are recruited leading to the activation of the TAK (TGF [transforming growth factor] β activated kinase)-TAB (TAK associated binding protein) complex. TAK is essential for the activation of the trimeric complex, IKK (inhibitor of IκB kinase), which in mammalian

cells is composed by IKKα, IKKβ and IKKγ. IKK activation leads to phosphorylation of the IκBα within an IκB/NFκB complex at amino acid residue serine 32 and serine 36. This phosphorylation of IκB is a marker for it to be tagged for ubiquitination. Once ubiquitinated it is degraded by the proteasome, thus releasing NFκB to translocate to the nucleus, where it binds to κB sites in the promoters and enhancers of its target genes.

Experimentalists have exposed a number of different cell types to NFκB activating stimuli, most widely done with TNFα and have observed nuclear-cytoplasmic oscillations of NFκB. Figure 6.2 shows time lapse fluorescent imaging of NFκB in SK-N-AS (human S-type neuroblastoma cells) cells after TNFα stimulation.

These nuclear-cytoplasmic oscillations are produced due to NFκB's activation of early gene products that are a source of negative feedback for NFκB. Such as its inhibitor IκBα and the IKK deactivator A20 (Skaug et al. 2011).

It is believed that the period of these oscillations serves as one of the mechanisms in determining which genes are later activated. That specific non-linear network motifs can decode frequencies.

6.3.1 Spatial stochastic model

We now present a spatial stochastic model of the canonical NFκB signalling pathway. Here when referring to NFκB we are specifically referring to RelA-NFκB1 heterodimer and IκB as IκBα. The computational domain consists of two concentric spheres and three individual gene sites. The outer sphere representing the cytoplasm and the inner sphere representing the nucleus.

The radius of the cytoplasm chosen to be $9.5\mu m$ and nucleus $5\mu m$, very close to that of HeLa cells. Later we might consider varying the nuclear and cytoplasmic volume in

order to analyse affects of extrinsic stochasticity. An unstructured mesh partitions the domain into the non intersecting subvolumes. For this initial model, NFκB is counted as a single molecule and we may add specific subunits and dimerisation in a later extension.

NFκB and IκB have a probability of binding to dynein proteins that transport them along microtubules toward the nucleus. The velocity field that the microtubule network gives rise to is modelled as a radially directed field and we take the rate of transport on the fibres, i.e the speed of the velocity field, to be $3 \times 10^{-5} m/min$ for the following simulations. Letting $\vec{(x)} = (x,y,z)$, be a point in our domain with Cartesian coordinates then the velocity field takes the form,

$$v\vec{(x)} = -\frac{v}{||\vec{x}||^2}(x,y,z)^T.$$

The density of the microtubules was calculated via the consideration that microtubules arise from the microtubule organising centre (MTOC), which resides close to the nucleus. The density of the microtubules are thus going to be denser near the nucleus than cellular membrane.

$$\rho(\vec{x}) = \rho_0 + (||\vec{x}|| - r_0)\frac{\rho_1 - \rho_0}{r_1 - r_0}, \quad \rho_0 = M/S_A^n, \rho_1 = M/S_A^c.$$

Where S_A^c is the surface area of the cellular membrane; r_1 is the radius of the cell; S_A^n is the surface area of the nuclear membrane and r_0 is the radius of the nucleus. We assume that the filament network consists of a total number of M microtubule filaments originating uniformly from the nuclear membrane and that the same number of filaments reach the cellular membrane. We choose M to be 150.

The promoter species are static. They are modelled in two states, changing only due

to two reactions. A free promoter becoming an occupied promoter due to the binding NFκB, or from occupied promoter to IκB bound occupied promoter. In the latter reaction, IκB can either induce degradation of the NFκB species and leave a free promoter and a free IκB molecule or it can capture NFκB in a IκB/NFκB complex and begin to transport it back to the cytoplasm. The three static species are confined to their localised gene sites and hence, given a diffusion coefficient of zero. Additionally, IKKa and A20 are restricted to diffusing within the cytoplasm only. All non zero cytoplasmic diffusion coefficients were given a value of $7 \times 10^{-11} m^2/min$ and a nuclear diffusion coefficient of $5 \times 10^{-11} m^2/min$. The initial state of the system was constructed with the copy number of all species being zero throughout the whole domain except the complex IκB/NFκB and the two free promoter sites pf_I and pf_A. One thousand IκB/NFκB species were randomly distributed in cytoplasmic sub-volumes at $t = 0$ and a single free promoter of IκB and A20 were put in their own specified voxel within the nucleus. The simulation ran for 1600 minutes. Due to the experimentally observed period of NFκB nuclear-cytoplasmic oscillations being around 100 minutes Nelson et al. (2004), WAVOS should be able to detect periods in and around this range for this amount of time.

The species involved in active transport are free IκB and free NFκB in the cytoplasm, directed radially inward toward the nuclear membrane.

The reactions used within the pathway follow that described in the tables below.

Table 6.1: *Cytoplasmic reactions*

Cytoplasmic Reaction	Description	Parameter value
$\text{Stimulus} \xrightarrow{\alpha} \text{IKKi}$	Production of inactive IKK via interaction with stimulus	$\alpha = 0.01825 M min^{-1}$
$\text{Stimulus} \xrightarrow{\alpha} \text{TAK}$	Productin of TAK via interaction with stimulus	$\alpha = 0.0185 M min^{-1}$
$\text{TAK} + \text{IKKi} \xrightarrow{A_1} \text{TAKIKKi}$	TAK and IKKi complex formation	$A_1 = 1 \times 10^{13} M^{-1} min^{-1}$
$\text{TAKIKKi} \xrightarrow{C_1} \text{TAK} + \text{IKKa}$	Activation of IKK	$C_1 = 50 M^{-1} min^{-1}$
$\text{IKKa} + \text{I}\kappa\text{BNF}\kappa\text{B} \xrightarrow{A_2} \text{IKKaI}\kappa\text{BNF}\kappa\text{B}$	Formation of IKKa and IκBNFκB complex	$A_2 = 2.5 \times 10^8 M^{-1} min^{-1}$
$\text{IKKaI}\kappa\text{BNF}\kappa\text{B} \xrightarrow{C_2} \text{IKKa} + \text{NF}\kappa\text{B}$	Catalytic degradation of IκB in the IKKaIκBNFκB complex	$C_2 = 25 M^{-1} min^{-1}$
$IkB + \text{NF}\kappa\text{B} \xrightarrow{\text{A}} \text{I}\kappa\text{BNF}\kappa\text{B}$	Formation of IκBNFκB complex	$A = 1 \times 10^{13} M^{-1} min^{-1}$
$\text{I}\kappa\text{Btran} \xrightarrow{\alpha_{pI}} \text{I}\kappa\text{Btran} + \text{I}\kappa\text{B}$	Translation of IκB protein	$\alpha_{pI} = 40 min^{-1}$
$\text{A20tran} \xrightarrow{\alpha_{pA}} \text{A20tran + A20}$	Translation of A20 protein	$\alpha_{pA} = 10 min^{-1}$
$\text{A20 + IKKa} \xrightarrow{\text{A}_3} \text{A20IKK}$	Formation of A20-IKKa complex	$A_3 = 1 \times 10^{11} M^{-1} min^{-1}$
$\text{A20IKKa} \xrightarrow{\text{C}_3} \text{A20 + IKKi}$	A20 induced deactivation of IKKa	$C_3 = 50 min^{-1}$
$\text{IKKi} \xrightarrow{\delta_S} \emptyset$	Degradation of IKKi	$\delta_S = 0.06 min^{-1}$
$A20 \xrightarrow{\delta_{pA}} \emptyset$	Degradation of A20	$\delta_{pA} = 0.07 min^{-1}$
$\text{NF}\kappa\text{B} \underset{\text{N}\sigma_{off}}{\overset{\text{N}\sigma_{on}}{\rightleftharpoons}} \text{NF}\kappa\text{Bmic}$	NFκB binding/unbinding to microtubule	$\text{N}\sigma_{on} = 1 \times 10^3 min^{-1}$ $\text{N}\sigma_{off} = 0.1 min^{-1}$
$\text{I}\kappa\text{B} \underset{\text{I}\sigma_{off}}{\overset{\text{I}\sigma_{on}}{\rightleftharpoons}} \text{I}\kappa\text{Bmic}$	Binding/unbinding of IκB to microtubule	$I\sigma_{on} = 1 \times 10^5 M^{-1} min^{-1}$ $I\sigma_{off} = 0.001 min^{-1}$
$\text{NF}\kappa\text{B}mic_i \xrightarrow{v} \text{NF}\kappa\text{B}mic_j$	Nuclear directed active transport along microtubules of NFκB between connected voxels	$v = 3 \times 10^{-5} mmin^{-1}$
$\text{I}\kappa\text{B}mic_i \xrightarrow{v} \text{I}\kappa\text{B}mic_j$	Nuclear directed active transport of IκB between connected voxels	$v = 3 \times 10^{-5} mmin^{-1}$
$\text{S}_{\text{ij}} \xrightarrow{\text{d}_{\text{jik}}} \text{S}_{\text{ik}}$	Molecular diffusion	$D = 6.5 \times 10^{-12} m^2 min^{-1}$

Table 6.2: *Reactions within IκB gene site*

IκB gene site Reaction	Description	Parameter value
$p_{fI} + \text{NF}\kappa\text{B} \underset{\text{k}_{\text{I2}}}{\overset{\text{k}_{\text{I1}}}{\rightleftharpoons}} p_{oI}$	Binding/unbinding of NFκB to the free IκB promoter	$k_{I1} = 7.35 \times 10^6 M^{-1} min^{-1}$ $k_{I2} = 0.06 min^{-1}$
$p_{oI} \xrightarrow{\alpha_I} p_{oI} + \text{I}\kappa\text{Btran}$	Induced transcription of IκB mRNA	$\alpha_I = 2.25 min^{-1}$
$p_{fI} \xrightarrow{\alpha_I/\gamma} p_{fI} + \text{I}\kappa\text{Btran}$	Basal transcription of IκB mRNA	$\alpha_I = 2.25 min^{-1}, \gamma = 1000$

Table 6.3: *Reactions within A20 gene site*

A20 gene site Reaction	Description	Parameter value
$p_{fA} + \text{NF}\kappa\text{B} \underset{\text{k}_{\text{A2}}}{\overset{\text{k}_{\text{A1}}}{\rightleftharpoons}} p_{oA}$	Binding/unbinding of NFκB to the free A20 promoter	$k_{A1} = 3.35 \times 10^6 M^{-1} min^{-1}$ $k_{A2} = 0.07 min^{-1}$
$p_{oA} \xrightarrow{\alpha_{mA}} p_{oA} + A20tran$	Induced transcription of A20 mRNA	$\alpha_{mA} = 2.25 min^{-1}$
$p_{fA} \xrightarrow{\alpha_{mA}/\gamma} p_{fA} + A20tran$	Basal transcription of A20 mRNA	$\alpha_{mA} = 2.25 min^{-1}, \gamma = 1000$

Table 6.4: *Global reactions*

Global Reactions	Description	Parameter value
$\mathrm{I}\kappa\mathrm{B} + \mathrm{NF}\kappa\mathrm{B} \xrightarrow{\mathrm{A_2}} \mathrm{I}\kappa\mathrm{BNF}\kappa\mathrm{B}$	Formation of IκB and NFκB complex	$A_2 = 2.5 \times 10^8 M^{-1} min^{-1}$
$\mathrm{I}\kappa\mathrm{Btran} \xrightarrow{d_m} \emptyset$	Degradation of IκBtran	$d_m = 0.0275 min^{-1}$
$\mathrm{I}\kappa\mathrm{B} \xrightarrow{d_p} \emptyset$	Degradation of IκB	$d_p = 0.07 min^{-1}$
$\mathrm{A20tran} \xrightarrow{d_m} \emptyset$	Degradation of A20tran	$d_m = 0.025 min^{-1}$
$\mathrm{A20} \xrightarrow{d_p} \emptyset$	Degradation of A20	$d_p = 0.07 min^{-1}$
Nuclear reactions	**Description**	**Parameter value used**
$IkB + \mathrm{NF}\kappa\mathrm{B} \xrightarrow{\mathrm{A}} \mathrm{I}\kappa\mathrm{BNF}\kappa\mathrm{B}$	Formation of IκBNFκB complex	$A = 1 \times 10^{13} M^{-1} min^{-1}$
$\mathrm{S_{ij}} \xrightarrow{d_{jik}} \mathrm{S_{ik}}$	Molecular diffusion of species	$Dm^2 min^{-1}$

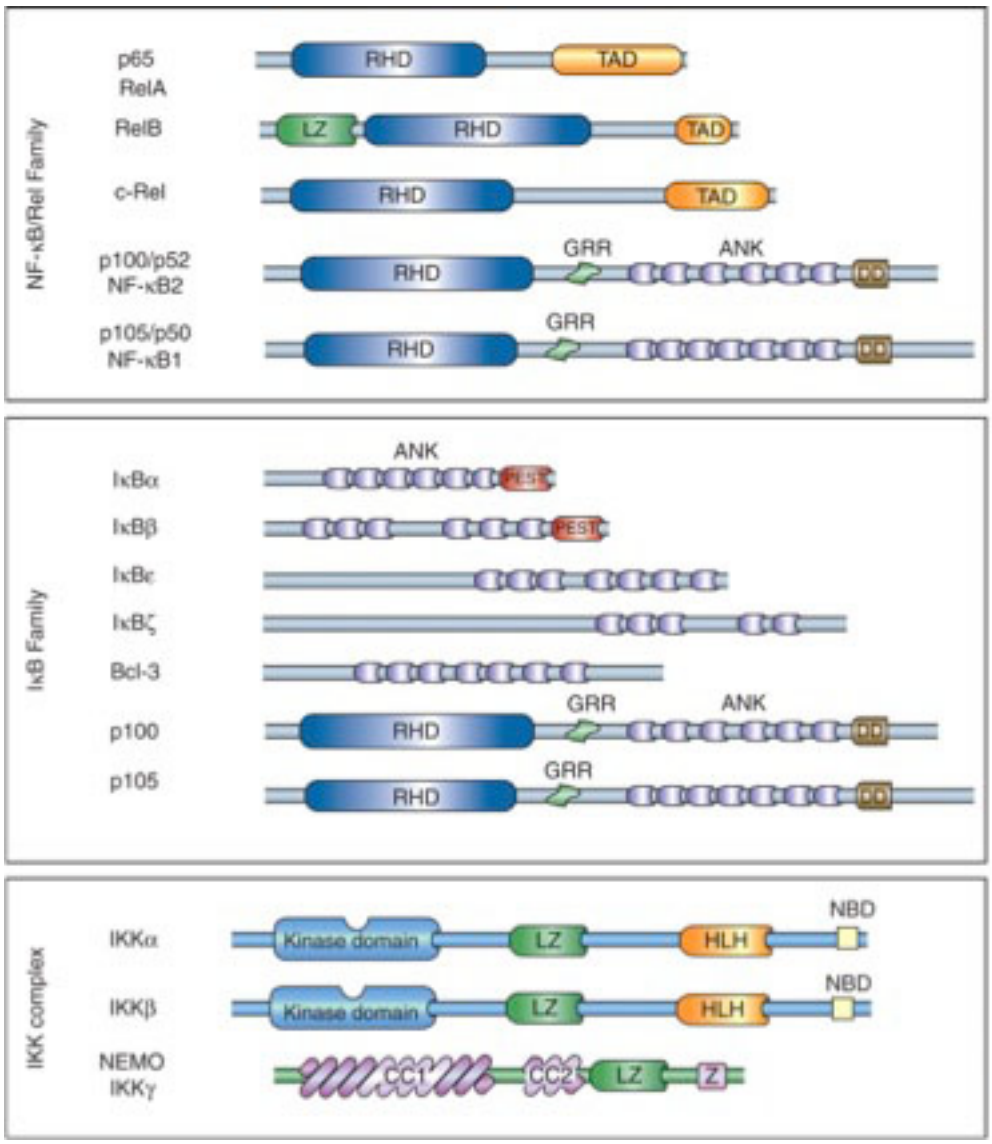

Figure 6.1: *The structural features of different protein families. Top: The family of NFκB subunits, Middle: The family of IκB proteins, Bottom: The family of IKK proteins. (RHD = Rel homology domain, TAD = Transcriptional activation domain, ANK = Ankyrin repeat, HLH = helix loop helix, LZ = Leucine zipper, GRR = Glycine rich region.) Hayden & Ghosh (2012)*

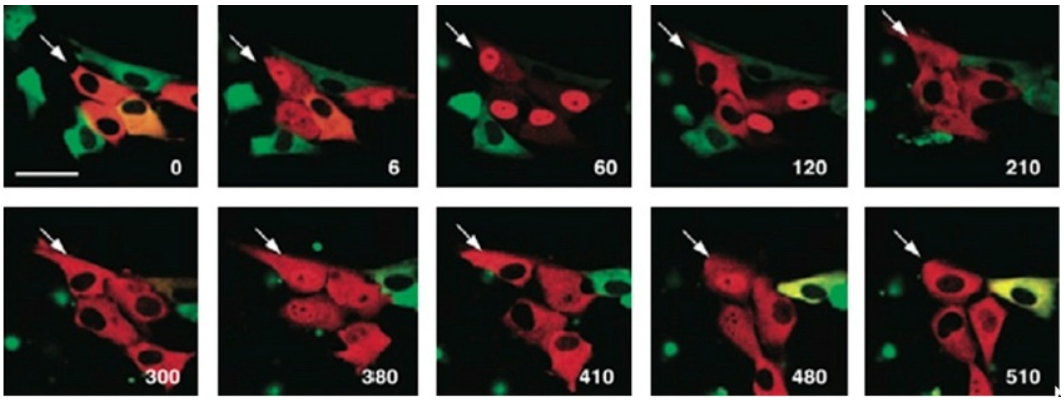

Figure 6.2: *Experimental observation of oscillations in NF-κB localisation. This figure shows time-lapse confocal images of NF-κB-containing species fused to a red fluorescent protein and of IκBa-containing species fused to a green fluorescent protein in SK-N-AS cells after stimulation with TNFa. The arrow marks one oscillating cell. Nuclear-cytoplasmic translocation of NF-κB-containing species is apparent. Time is shown in minutes and the scale bar represents 50μm.*

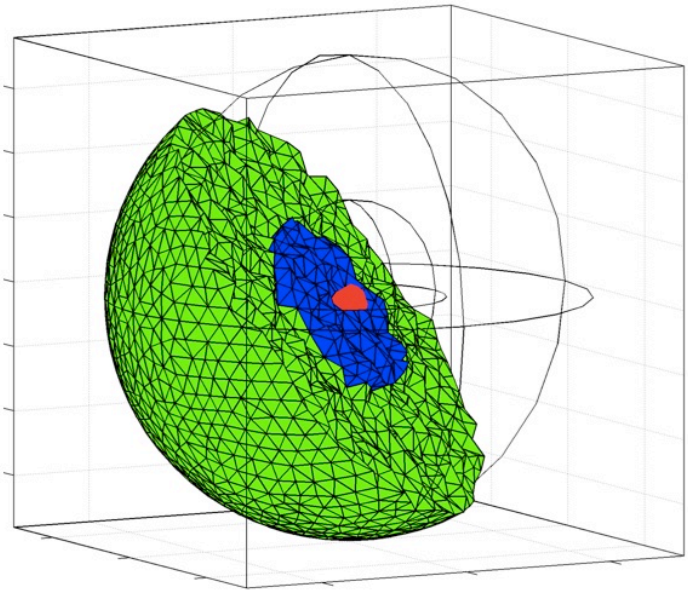

Figure 6.3: *Schematic cell*

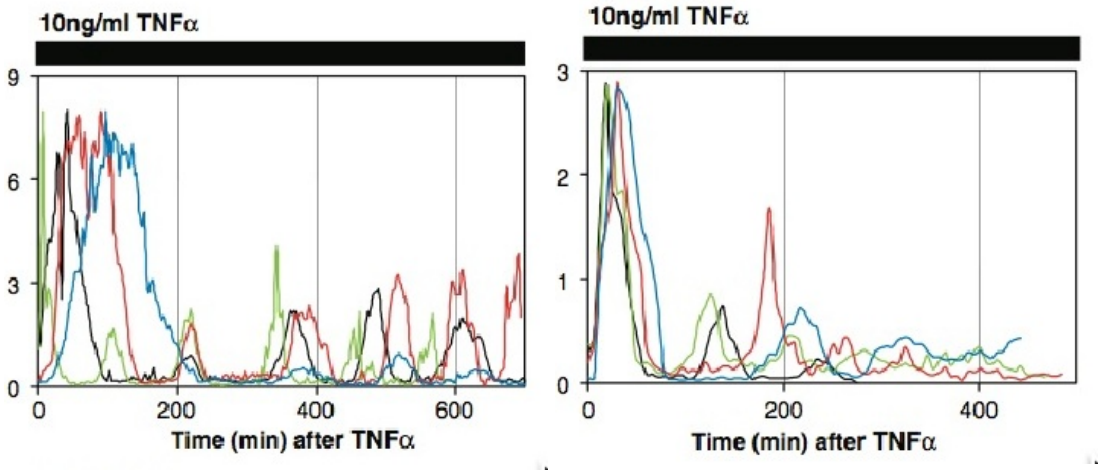

Figure 6.4: *Analysis of the dynamics of NF-κB localization and κB dependent reporter gene expression. Time course of N:C localization of RelA-DsRed in cells co-expressing IκB-EGFP. N:C ratio in RelA-DsRed fluorescence was normalized to highest peak intensity. The peak N:C ratio was expressed as the average value for each set of four cells. Data from each cell is represented by a different colored line. Left: SK-N-AS cells treated with continual 10 ng/ml TNF. Right: HeLa cells treated with continual 10 ng/ml TNF.*

6.3.2 Gene sites clustered at the centre of the nucleus

We now consider our spatial-stochastic model of the NFκB GRN with gene sites clustered at the centre of the nucleus. We then position the gene site for IκBα at $(x_{I\kappa B}, y_{I\kappa B}, z_{I\kappa B}) = (0.5\mu m, 0\mu m, 0\mu m)$ and the gene for A20 at $(x_{A20}, y_{A20}, z_{A20}) = (-0.5\mu m, 0\mu m, 0\mu m)$.

For the simulations, to best compare to experimental data, we pay most attention to the nuclear-cytoplasmic ratio, oscillations of NF-κB within the nucleus and the spatial snapshots.

We now present simulations, with gene sites clustered around the centre of the nucleus. For each individual simulation we present two figures: a time series of each species and a series of ten spatial snapshots of the distribution of NFκB throughout the cell at specific times chosen to reflect the highest peaks of NFκB levels in the nucleus and cytoplasm. Each snapshot is of a slice of the cellular domain, showing the x-y plane. We investigate how changing the diffusion parameter D changes the spatial-temporal dynamics of the NFκB GRN. We present two simulation results for each parameter D chosen. First, we present two simulations for the smaller diffusion parameter $D = 1 \times 10^{-12}$ and two each as we increase D. We D increase from $D = 1 \times 10^{-12}$ to $D = 8 \times 10^{-12}$ to finally $D = 3 \times 10^{-11}$.

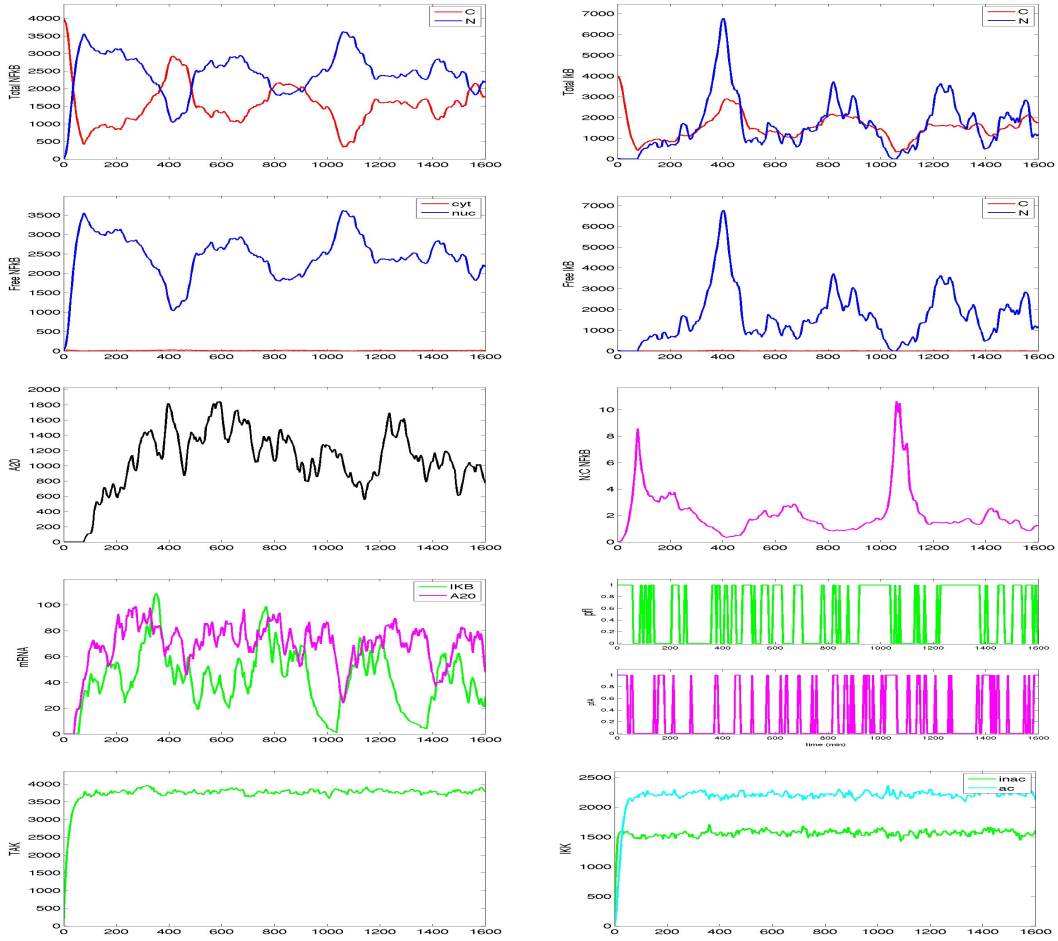

Figure 6.5: *Time series of species in the NFκB pathway. For each species, with the exception of NFκB and IκBα (in which we show nuclear and cytoplasmic numbers) the plots show the total copy number over the entire cell against time. Gene sites are clustered close to the centre of the nucleus and* $D = 1 \times 10^{-12}$. *From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, Free IκB, A20, nuclear-cytoplasmic ration of NFκB, mRNA of IκB and A20, Free promoter species of IκB and A20, TAK, IKK (inactive and active).*

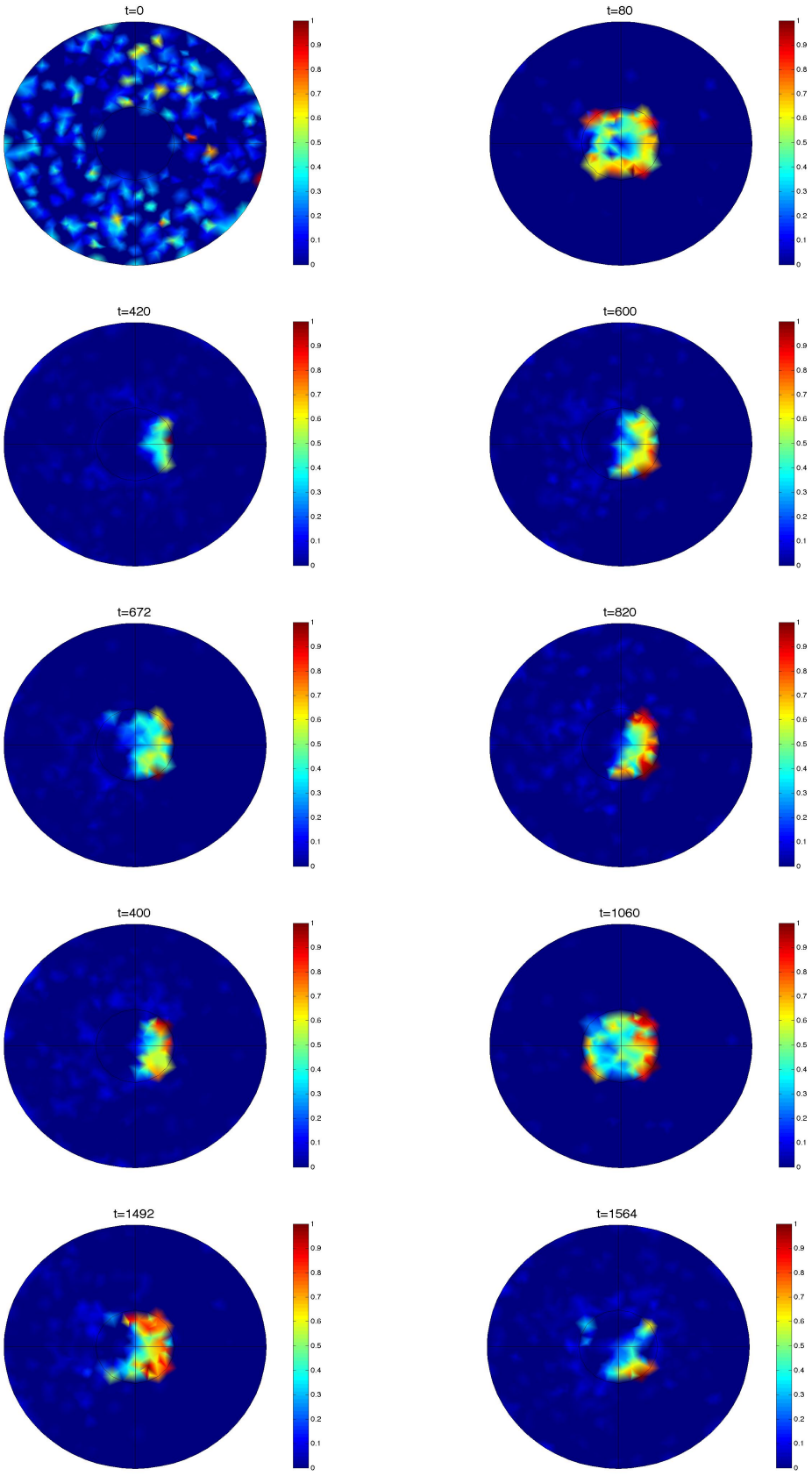

Figure 6.6: *Time series of species in the NFκB pathway. Gene sites are clustered at the origin and* $D = 1 \times 10^{-12}$. *From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, Free IκB, A20, nuclear-cytoplasmic ration of NFκB, mRNA of IκB and A20, Free promoter species of IκB and A20, TAK, IKK (inactive and active).*

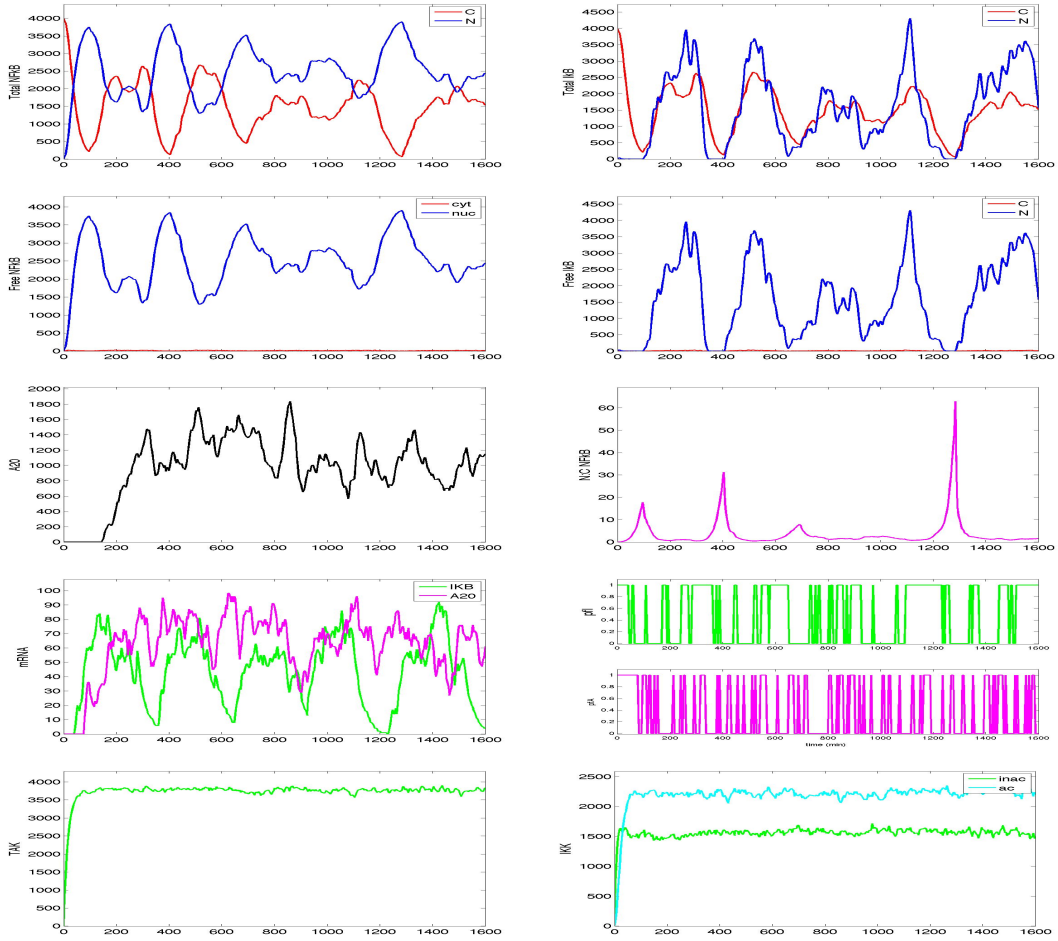

Figure 6.7: *Time series of species in the NFκB pathway. For each species, with the exception of NFκB and IκBα (in which we show nuclear and cytoplasmic numbers) the plots show the total copy number over the entire cell against time. Gene sites are clustered close to the centre of the nucleus and* $D = 1 \times 10^{-12}$*. From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, Free IκB, A20, nuclear-cytoplasmic ration of NFκB, mRNA of IκB and A20, Free promoter species of IκB and A20, TAK, IKK (inactive and active).*

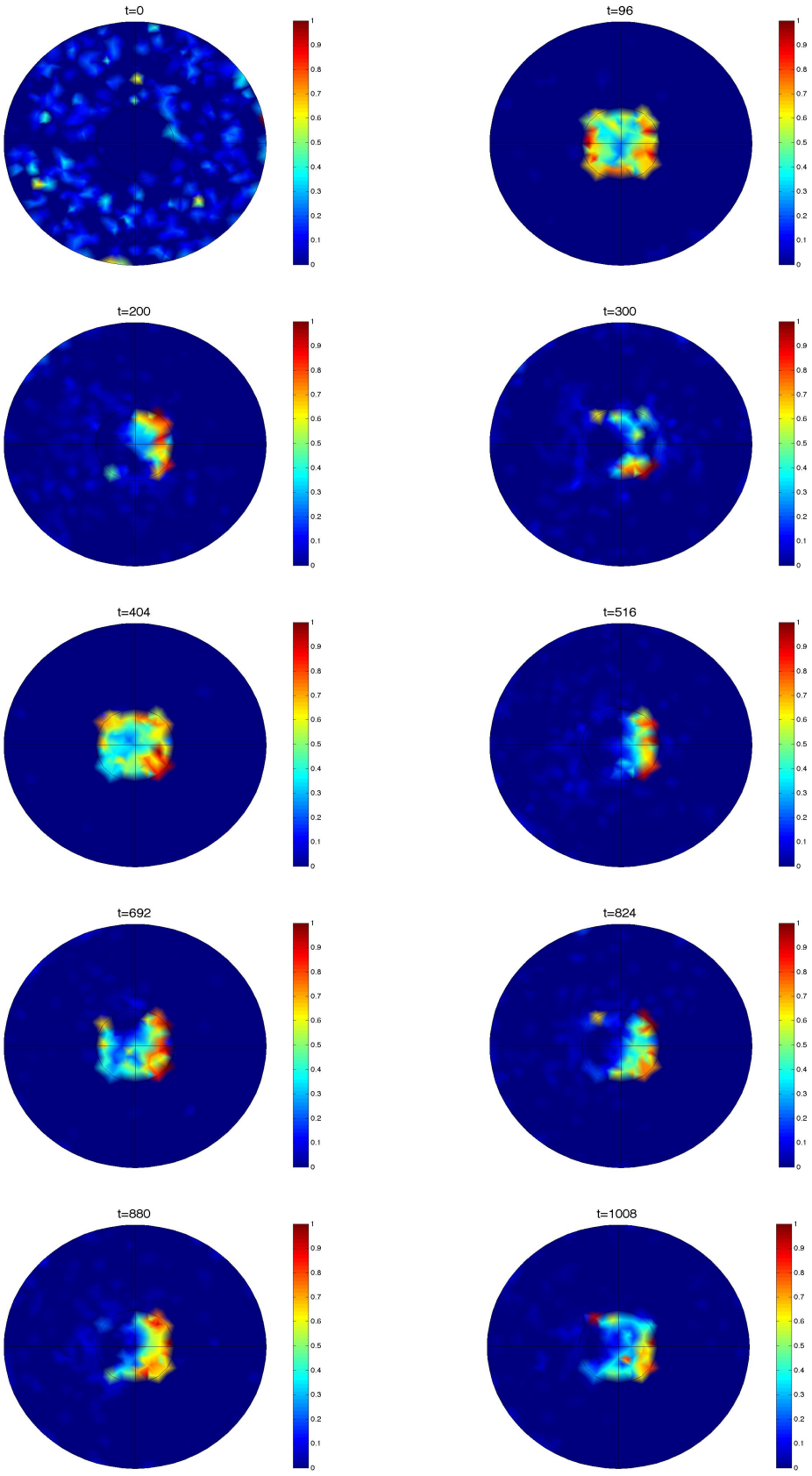

Figure 6.8: *Time series of species in the NFκB pathway. Gene sites are clustered at the origin and* $D = 1 \times 10^{-12}$. *From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, Free IκB, A20, nuclear-cytoplasmic ration of NFκB, mRNA of IκB and A20, Free promoter species of IκB and A20, TAK, IKK (IKKi and IKKa).*

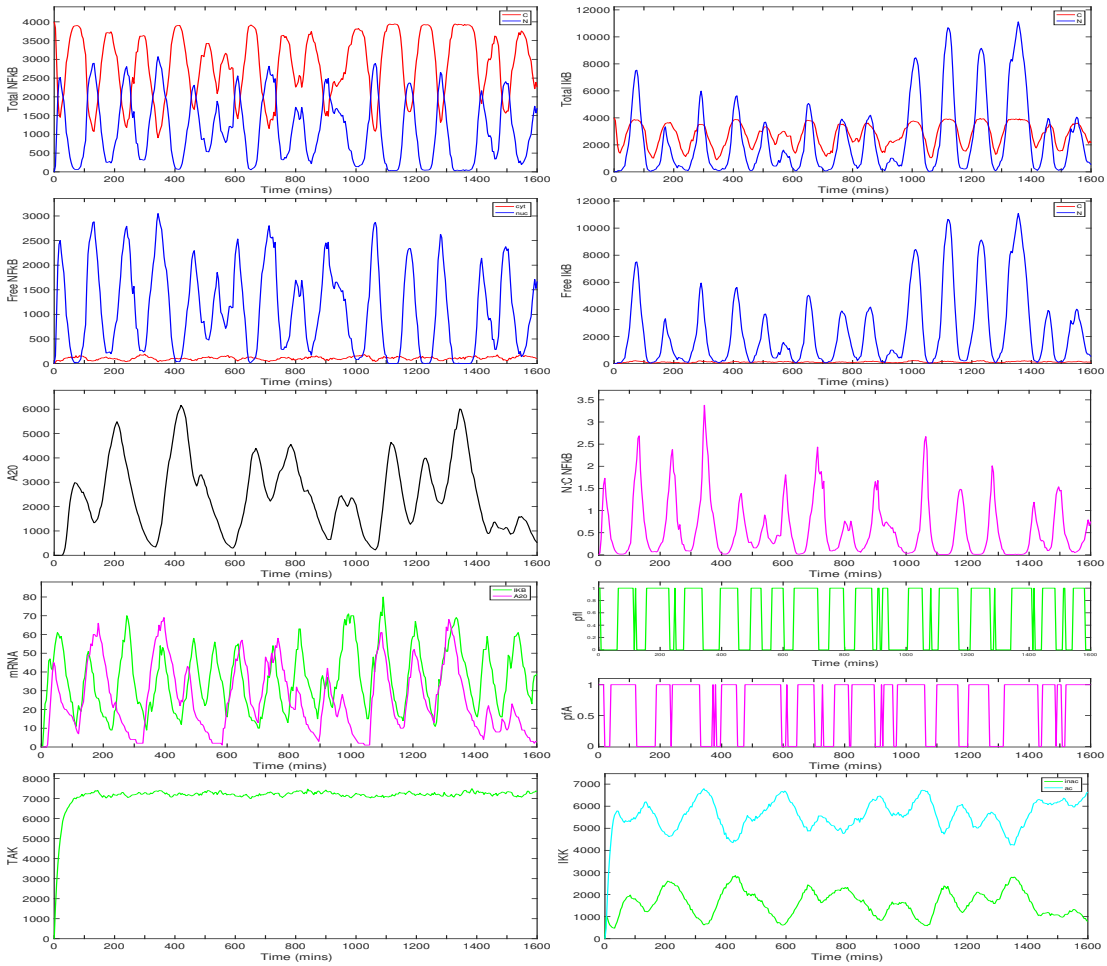

Figure 6.9: *Time series of species in the NFκB pathway. For each species, with the exception of NFκB and IκBα (in which we show nuclear and cytoplasmic numbers) the plots show the total copy number over the entire cell against time. Gene sites are clustered close to the centre of the nucleus and* $D = 8 \times 10^{-12}$*. From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, Free IκB, A20, nuclear-cytoplasmic ration of NFκB, mRNA of IκB and A20, Free promoter species of IκB and A20, TAK, IKK (inactive and active).*

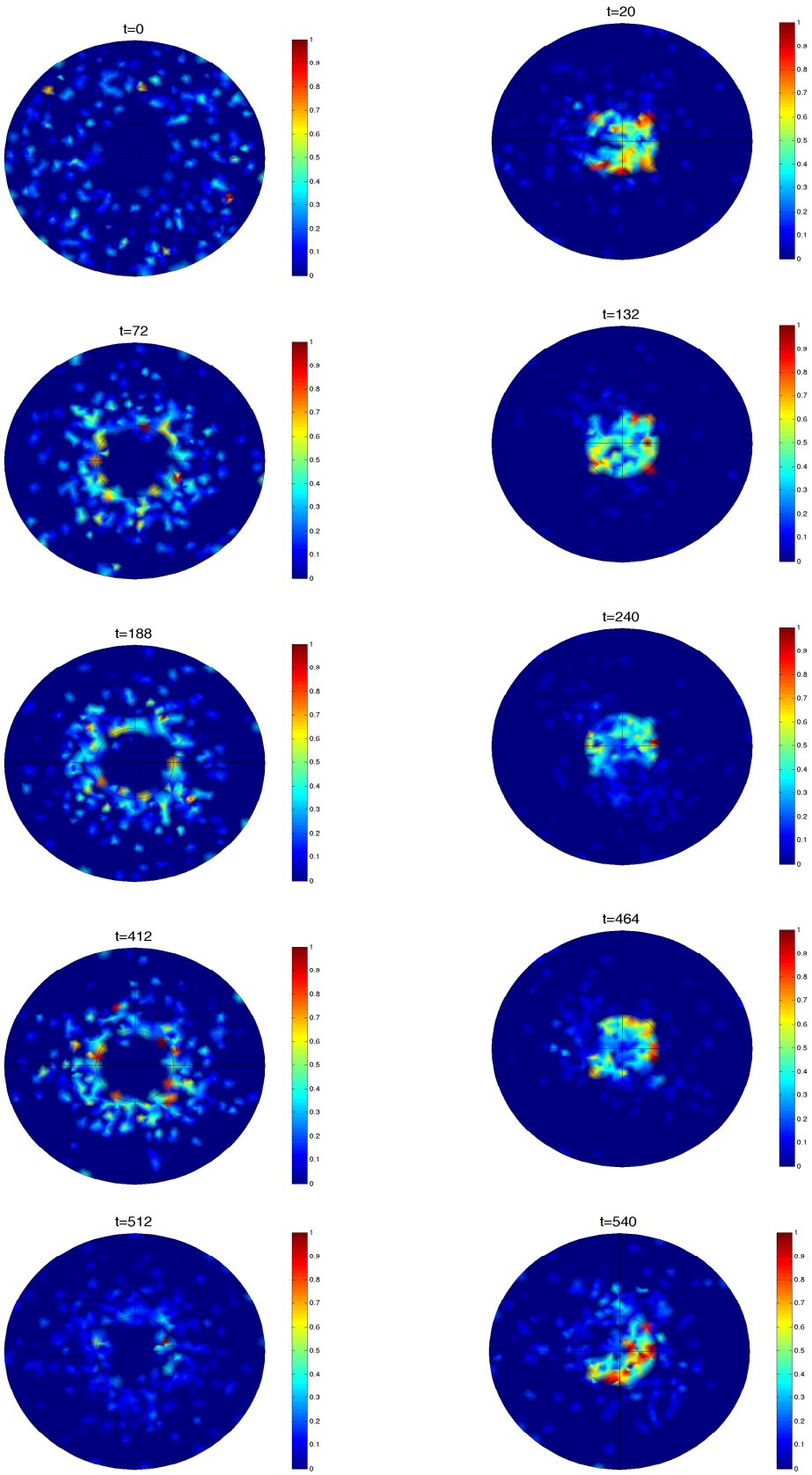

Figure 6.10: *Spatial snapshots of NFκB. Gene sites are clustered at the origin and* $D = 8 \times 10^{-12}$

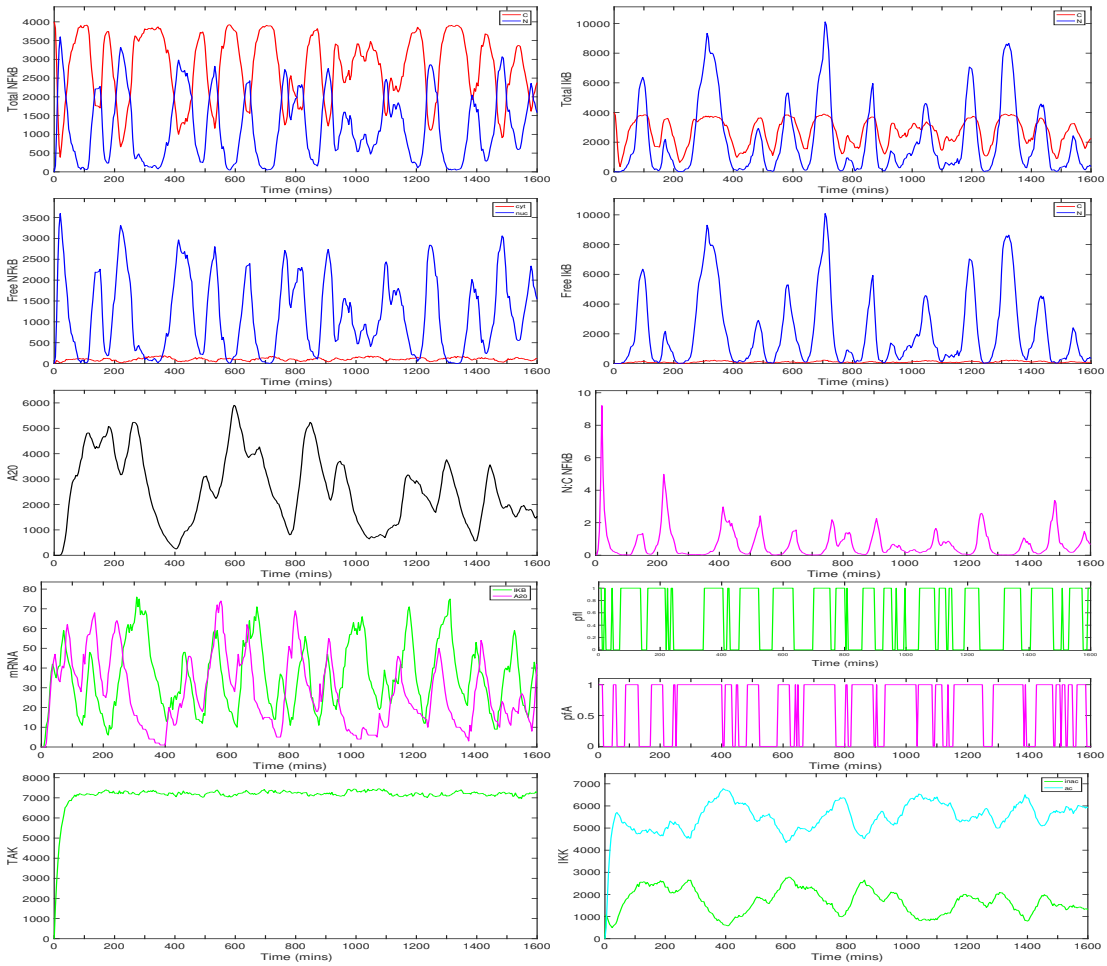

Figure 6.11: *Time series of species in the NFκB pathway. For each species, with the exception of NFκB and IκBα (in which we show nuclear and cytoplasmic numbers) the plots show the total copy number over the entire cell against time. Gene sites are clustered close to the centre of the nucleus and* $D = 8 \times 10^{-12}$. *From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, Free IκB, A20, nuclear-cytoplasmic ration of NFκB, mRNA of IκB and A20, Free promoter species of IκB and A20, TAK, IKK (inactive and active).*

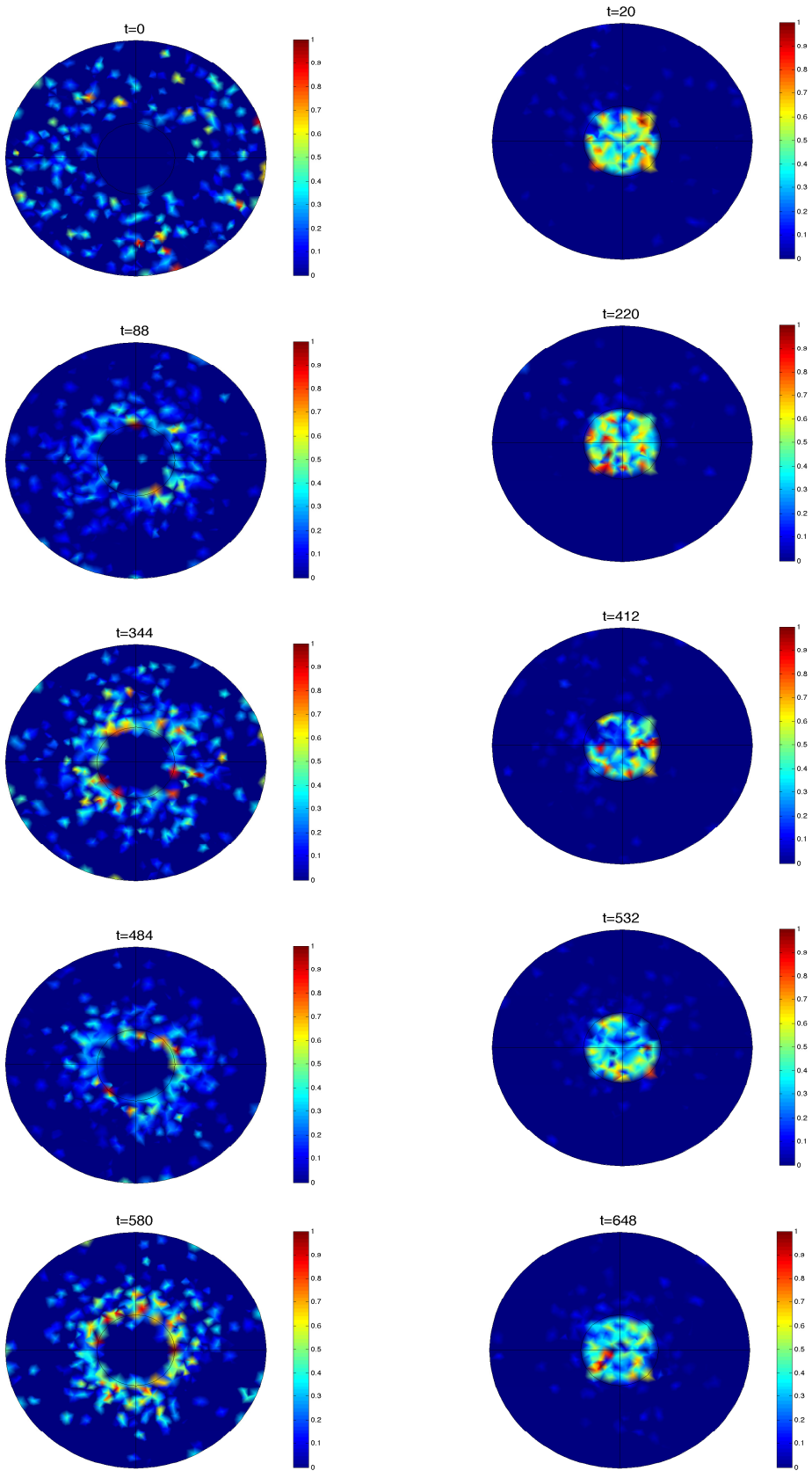

Figure 6.12: *Spatial snapshots of NFκB. Gene sites are clustered at the origin and* $D = 8 \times 10^{-12}$

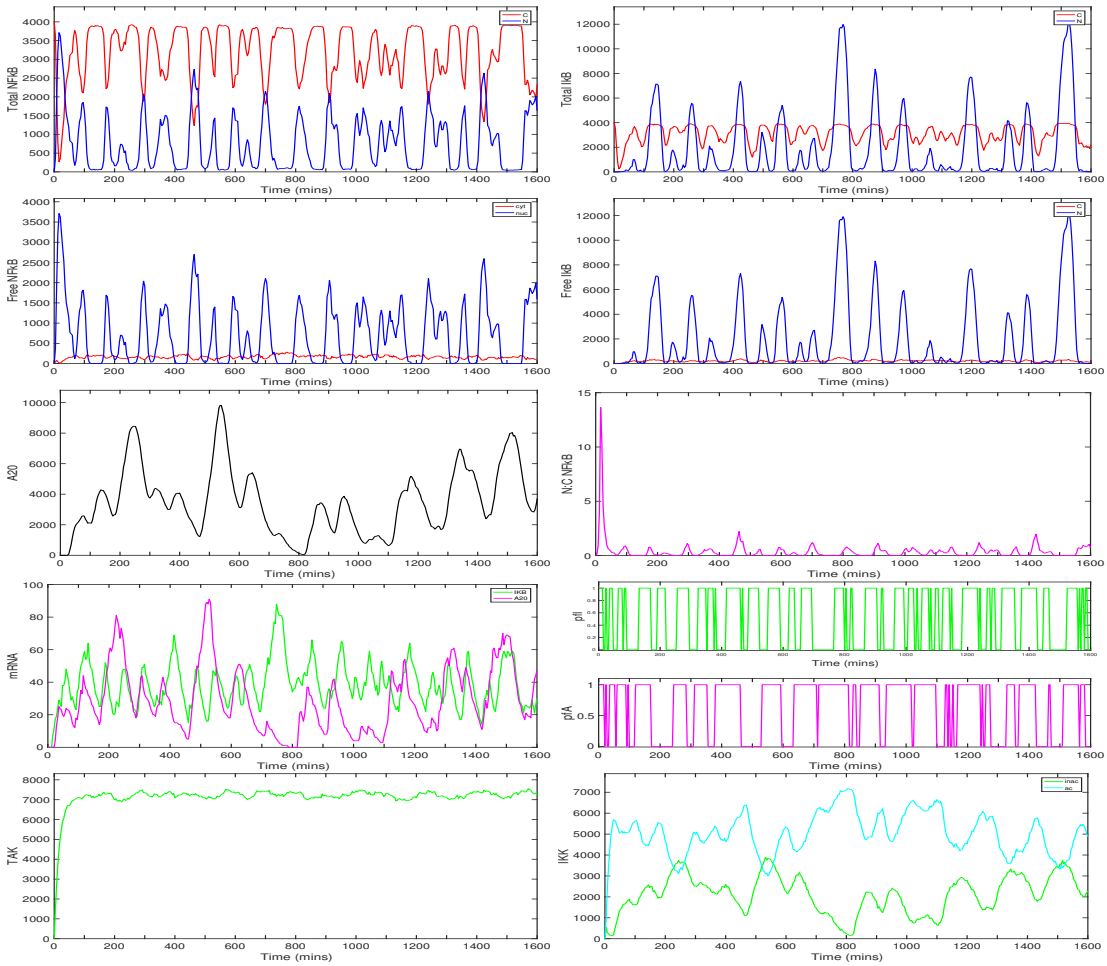

Figure 6.13: *Time series of species in the NFκB pathway. For each species, with the exception of NFκB and IκBα (in which we show nuclear and cytoplasmic numbers) the plots show the total copy number over the entire cell against time. Gene sites are clustered close to the centre of the nucleus and* $D = 3 \times 10^{-11}$*. From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, Free IκB, A20, nuclear-cytoplasmic ration of NFκB, mRNA of IκB and A20, Free promoter species of IκB and A20, TAK, IKK (inactive and active).*

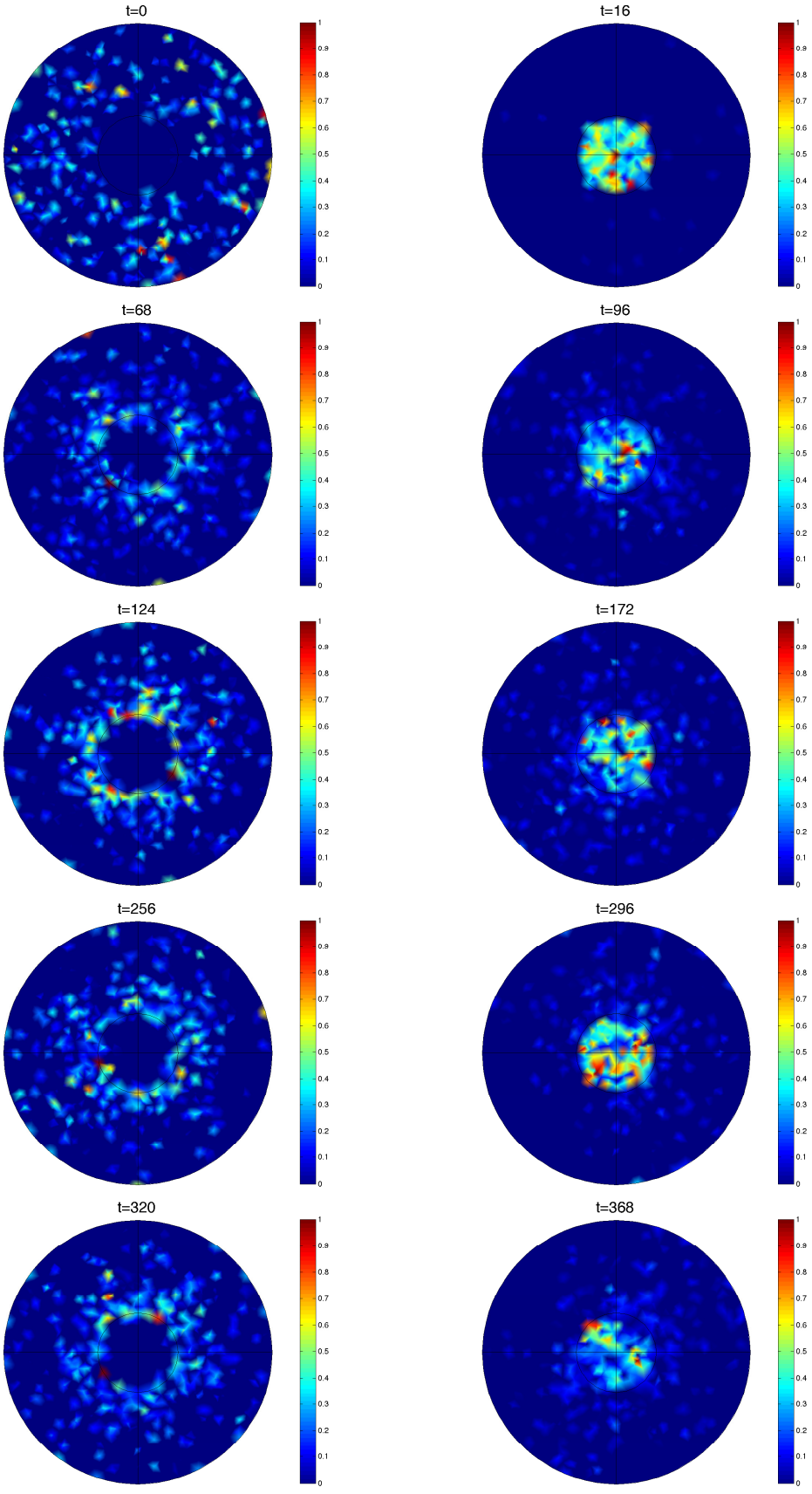

Figure 6.14: *Snapshots of NFκB. Gene sites clustered clustered at the centre of the nucleus.* $D = 3 \times 10^{-11}$

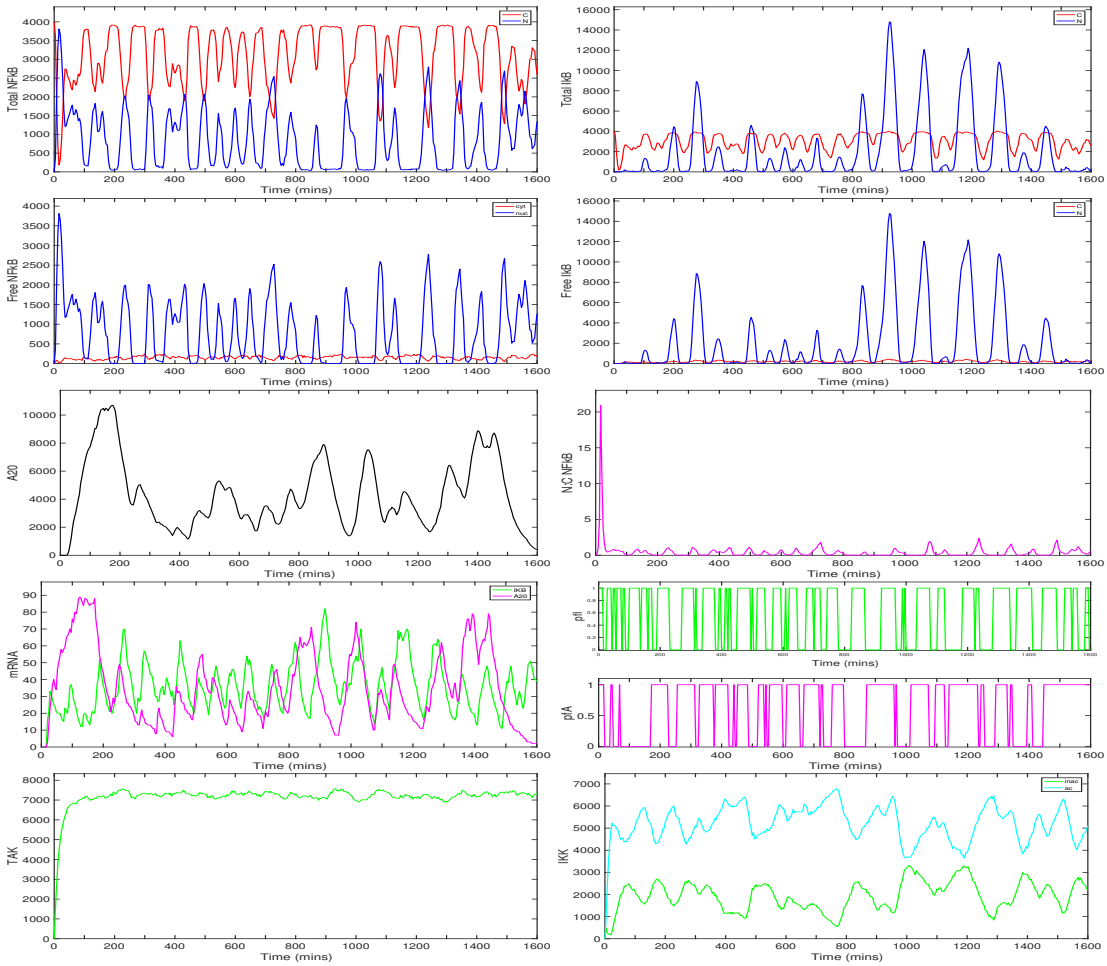

Figure 6.15: *Time series of species in the NFκB pathway. For each species, with the exception of NFκB and IκBα (in which we show nuclear and cytoplasmic numbers) the plots show the total copy number over the entire cell against time. Gene sites are clustered close to the centre of the nucleus and* $D = 3 \times 10^{-11}$. *From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, Free IκB, A20, nuclear-cytoplasmic ration of NFκB, mRNA of IκB and A20, Free promoter species of IκB and A20, TAK, IKK (inactive and active).*

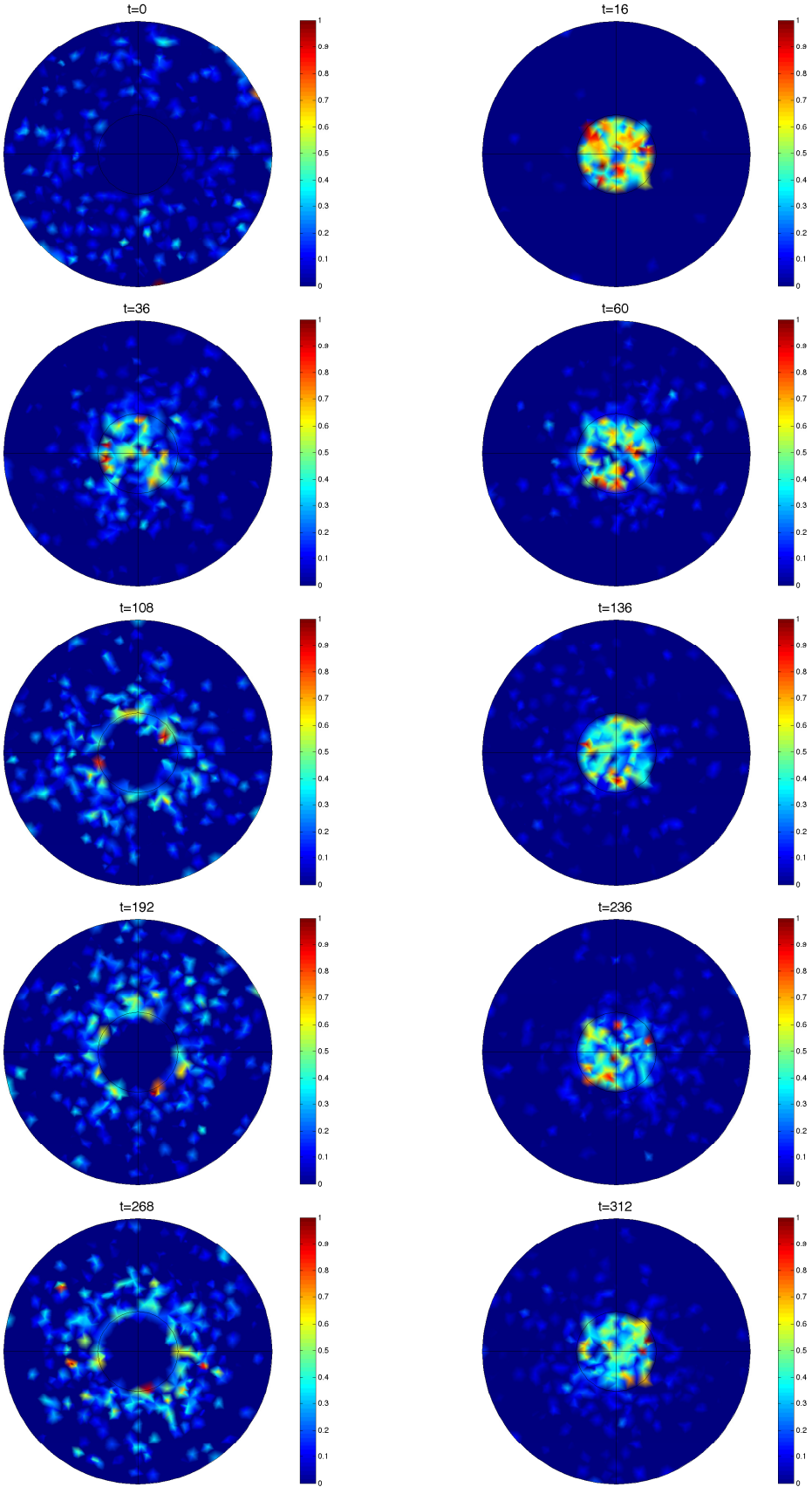

Figure 6.16: *Snapshots of NFκB. Gene sites clustered clustered at the centre of the nucleus.* $D = 3 \times 10^{-11}$

In the smaller diffusion regime $D = 1 \times 10^{-12}$ there are long periods seen of nuclear NFκB oscillations, see the top left images of figures 6.5 and 6.7. The oscillations seen are also of relatively low amplitude and so do not exhibit much influence on the nuclear-cytoplasmic oscillatory levels of NFκB. We deem this slower diffusion regime as being incapable of producing the nuclear-cytoplasmic oscillations seen in the experimental realm. NFκB tends to accumulate more in the nucleus for this slower diffusion regime. This is due to the fact that the transcribed IκB mRNA, induced by the NFκB-IκB promoter complex, takes longer to reach the cytoplasm. Thus, less IκB protein can be synthesised and hence, there is less down-regulation of NFκB i.e. less sequestration of NFκB by IκB in the cytoplasm.

For the middle diffusion regime, $D = 8 \times 10^{-12}$, somewhat steady, nuclear-cytoplasmic oscillations occur with a somewhat regular period and amplitude, see figures, 6.9, 6.10, 6.11 and 6.12. The period of nuclear NFκB oscillations varies between 83 minutes and 145 minutes. The average period seen was 108 minutes. For this diffusion regime, there was a somewhat steady level of NFκB-IκB and NFκB-A20 promoter binding and unbinding. Due to the faster diffusion regime, IκB mRNA and A20 mRNA, (which can be transcribed due to their occupied promoters) can readily reach the cytoplasm where IκB and A20 protein can then be synthesised at a higher rate, due to the higher number of mRNA that can reach the cytoplasm. Thus, this creates a higher probability in down-regulation of NFκB levels. IκB can reach the nucleus faster to form the inhibiting complex, IκB-NFκB and sequester NFκB to the cytoplasm and A20 can interact with the active form of IKK, (IKKa) and inactivate it to form IKKi. Thus, the levels of NFκB decrease. The whole process can then begin again. For this diffusion regime, we see this cycle occur very regularly.

For the faster diffusion regime, $D = 3 \times 10^{-11}$, oscillations also occur at a somewhat regular frequency. Although this time they increase in frequency. Here diffusion aids in a greater feedback, due to the even higher level of mRNA being able to reach the

cytoplasm and hence for IκB and A20 to be synthesised at a higher rate also. Due to the fact that NFκB can be transported in the cell quicker, it also reaches its target genes faster. This is reflected in the greater level of promoter activity, see the promoter species in figures 6.13 and 6.15. However, a greater increase in IκB and A20 leads to a greater increase in the down-regulation of NFκB. This resulted in the maximum copy number of nuclear NFκB molecules decreasing. The period of nuclear NFκB oscillations varies between 66 minutes and 133 minutes. The average period seen was 83 minutes. The period observed appeared to look rougher than for the middle diffusion regime. This rougher period was also seen in the levels of A20 compared to the, $D = 8 \times 10^{-12}$ diffusion regime.

For the middle and faster diffusion regime we see oscillations reminiscent of those captured by fluorescent imaging techniques, see figures 6.10, 6.12, 6.14 and 6.16 to compare with figure 6.2 and 6.4.

Common to all diffusion regimes was the appearance of an initial spike in nuclear NFκB.

6.3.3 Gene sites clustered close to the nuclear membrane

We now present simulations, with gene sites clustered close to the nucleus membrane. We then position the gene site for IκBα at $(x_{I\kappa B}, y_{I\kappa B}, z_{I\kappa B}) = (0.5\mu m, 2.85\mu m, 0\mu m)$ and the gene for A20 at $(x_{A20}, y_{A20}, z_{A20}) = (-0.5\mu m, 2.85\mu m, 0\mu m)$. For each individual simulation we present two figures: a time series of each species and a series of ten spatial snapshots of the distribution of NFκB throughout the cell at specific times chosen to reflect the highest peaks of NFκB levels in the nucleus and cytoplasm. Each snapshot is of a slice of the cellular domain, showing the x-y plane. We investigate how changing the diffusion parameter D changes the spatial-temporal dynamics of the

NFκB GRN for the cluster of gene sites close to the nucleus membrane. We present two simulation results for each parameter D chosen. First, we present two simulations for the smaller diffusion parameter $D = 8 \times 10^{-12}$ and two for each increase in D. We increase D from $D = 8 \times 10^{-12}$ to $D = 3 \times 10^{-11}$ to finally $D = 1 \times 10^{-10}$. For the smaller diffusion parameter, $D = 1 \times 10^{-12}$ that we used for the gene site cluster at the centre of the nucleus there were no oscillations seen within the oscillatory range of experimental data, so we ignore this value of D.

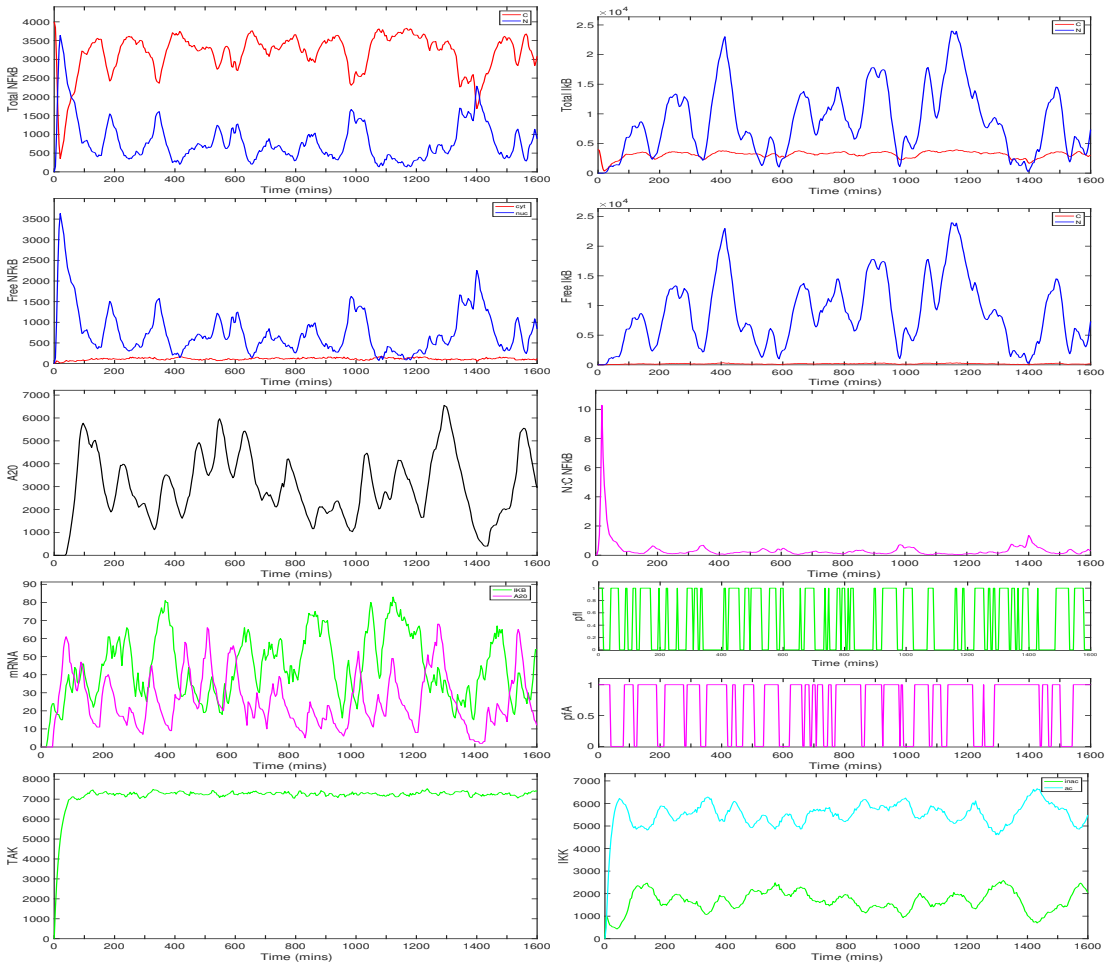

Figure 6.17: *Time series of species in the NFκB pathway. For each species, with the exception of NFκB and IκBα (in which we show nuclear and cytoplasmic numbers) the plots show the total copy number over the entire cell against time. Gene sites are clustered close to the nuclear membrane and* $D = 8 \times 10^{-12}$. *From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, Free IκB, A20, nuclear-cytoplasmic ration of NFκB, mRNA of IκB and A20, Free promoter species of IκB and A20, TAK, IKK (inactive and active).*

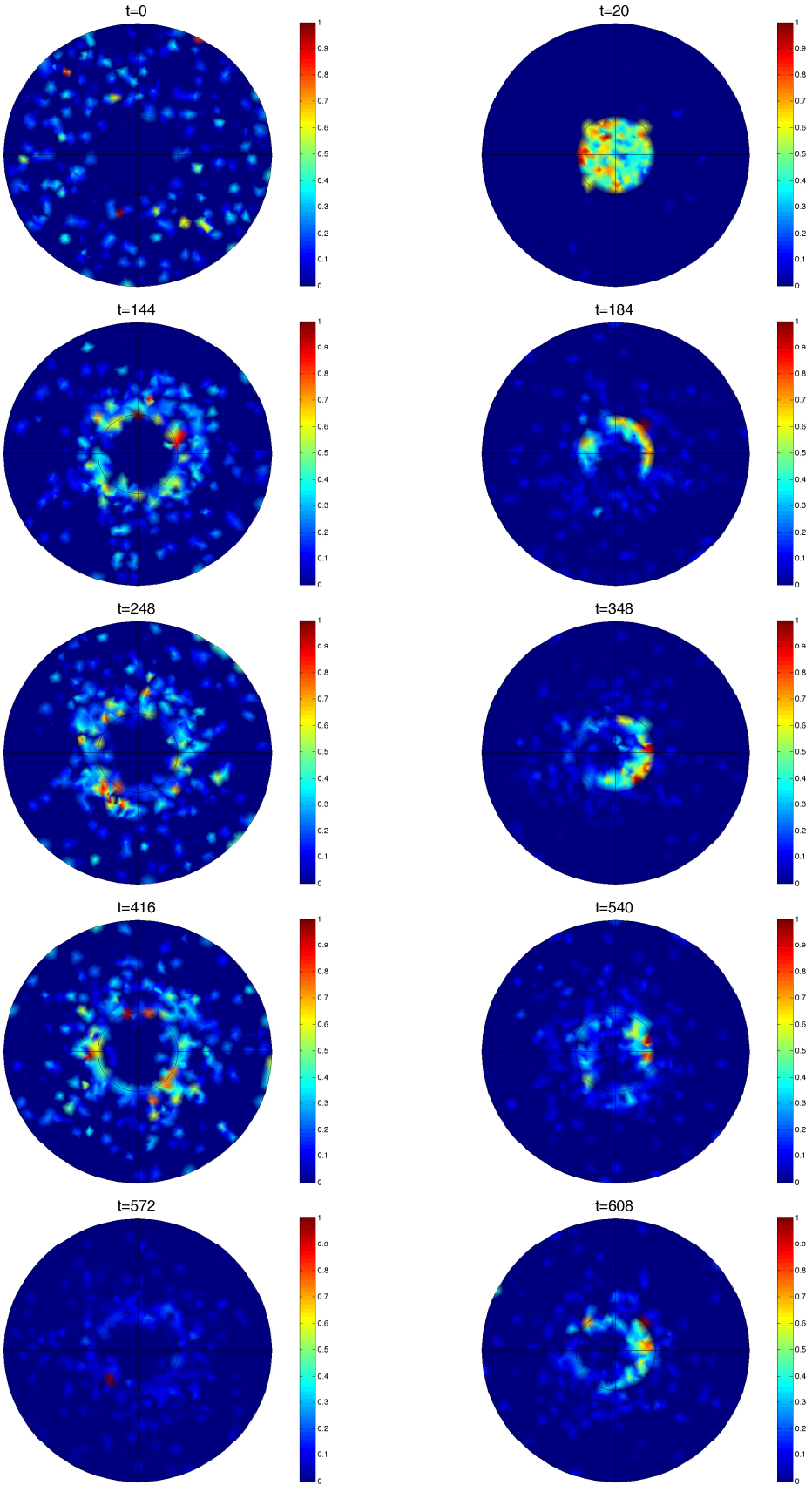

Figure 6.18: *Snapshots of NFκB. Gene sites clustered close to the nuclear membrane.* $D = 8 \times 10^{-12}$

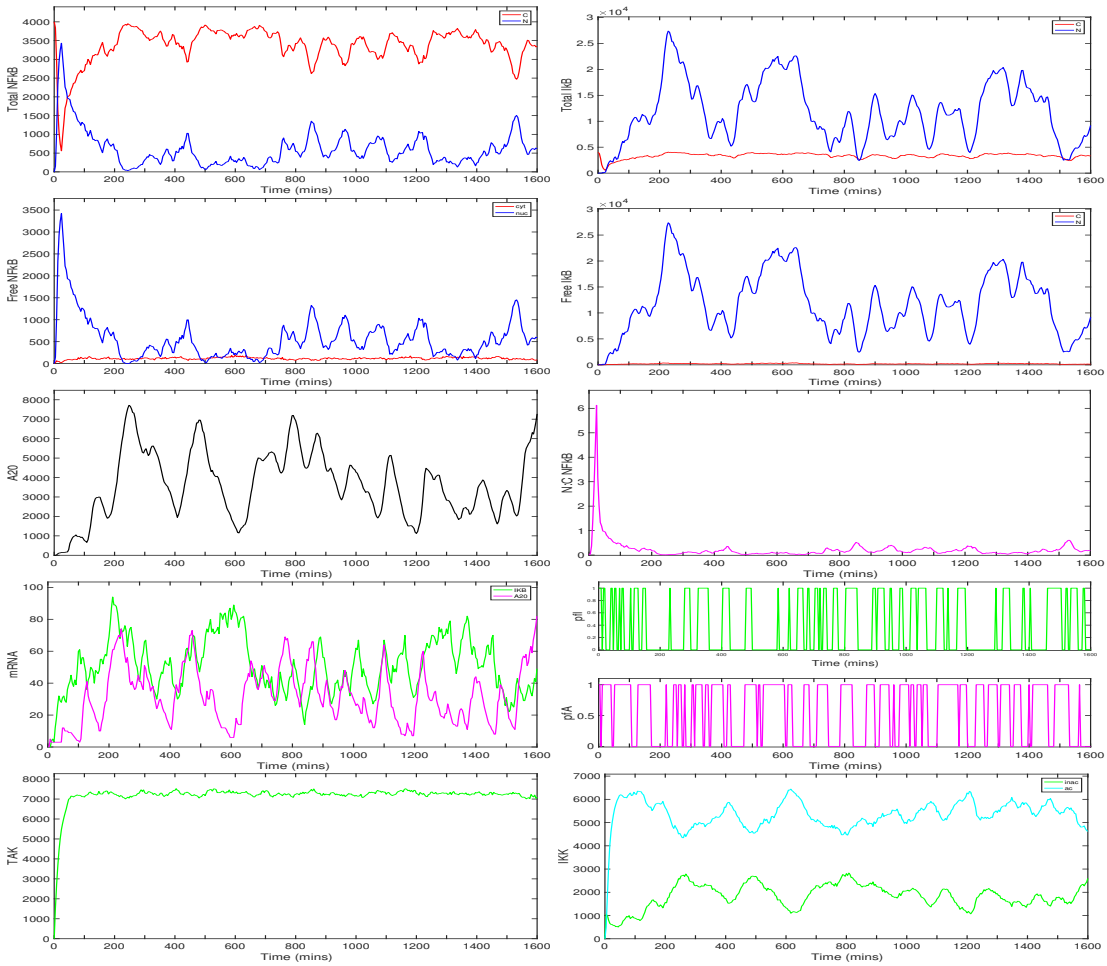

Figure 6.19: *Time series of species in the NFκB pathway. For each species, with the exception of NFκB and IκBα (in which we show nuclear and cytoplasmic numbers) the plots show the total copy number over the entire cell against time. Gene sites are clustered close to the nuclear membrane and* $D = 8 \times 10^{-12}$. *From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, Free IκB, A20, nuclear-cytoplasmic ration of NFκB, mRNA of IκB and A20, Free promoter species of IκB and A20, TAK, IKK (inactive and active).*

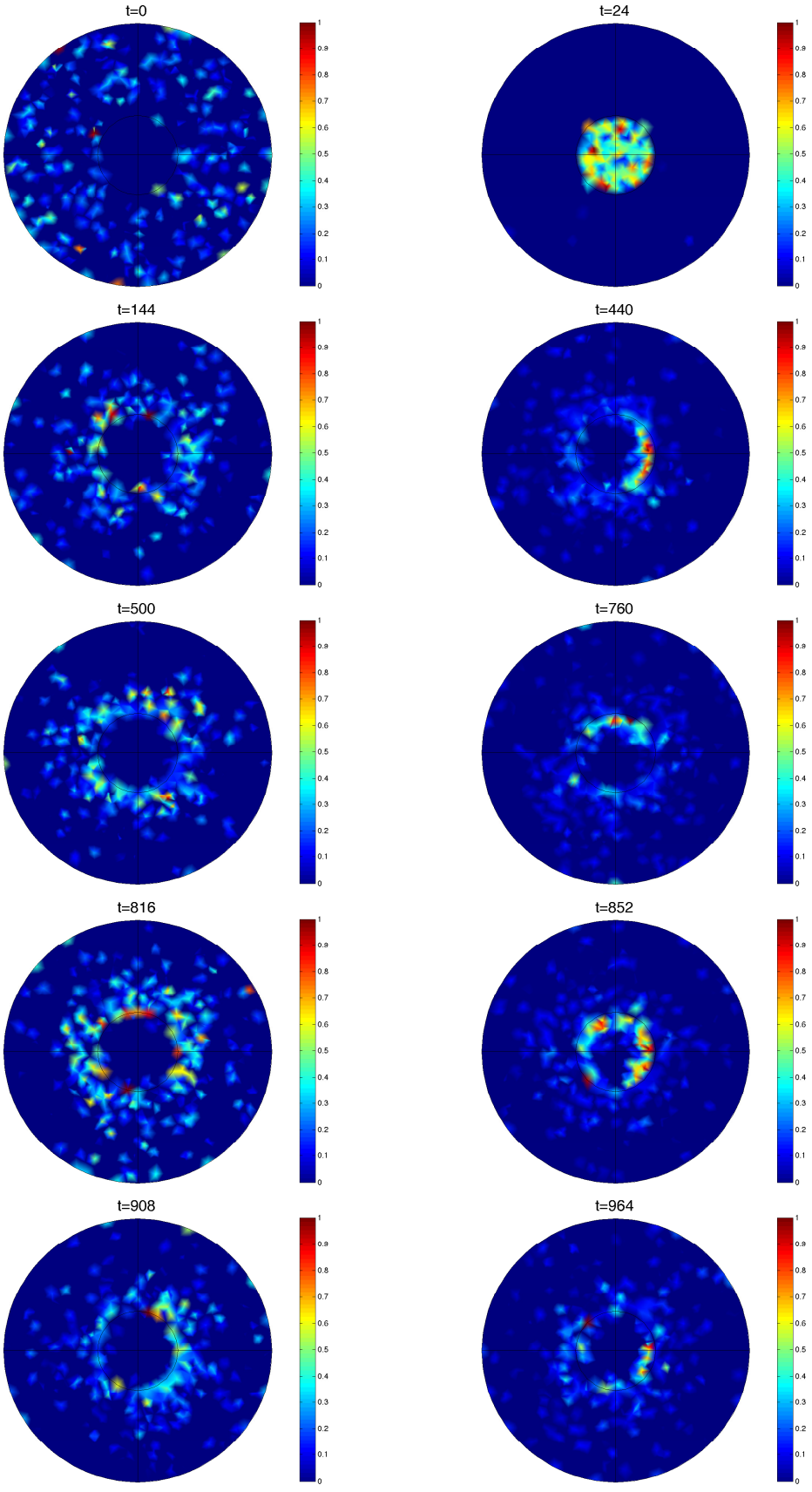

Figure 6.20: *Snapshots of NFκB. Gene sites clustered close to the nuclear membrane.* $D = 8 \times 10^{-12}$

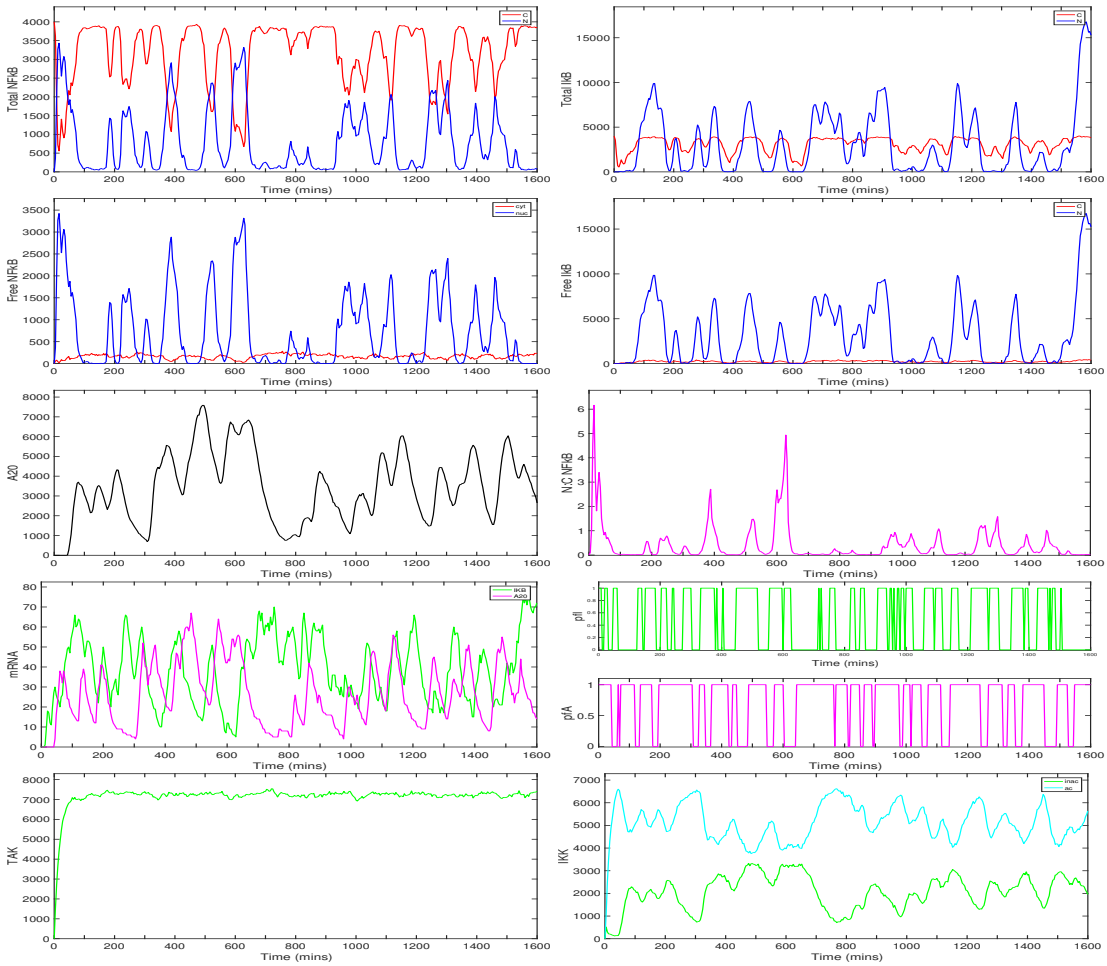

Figure 6.21: *Time series of species in the NFκB pathway. For each species, with the exception of NFκB and IκBα (in which we show nuclear and cytoplasmic numbers) the plots show the total copy number over the entire cell against time. Gene sites are clustered close to the nuclear membrane and* $D = 3 \times 10^{-11}$. *From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, Free IκB, A20, nuclear-cytoplasmic ration of NFκB, mRNA of IκB and A20, Free promoter species of IκB and A20, TAK, IKK (inactive and active).*

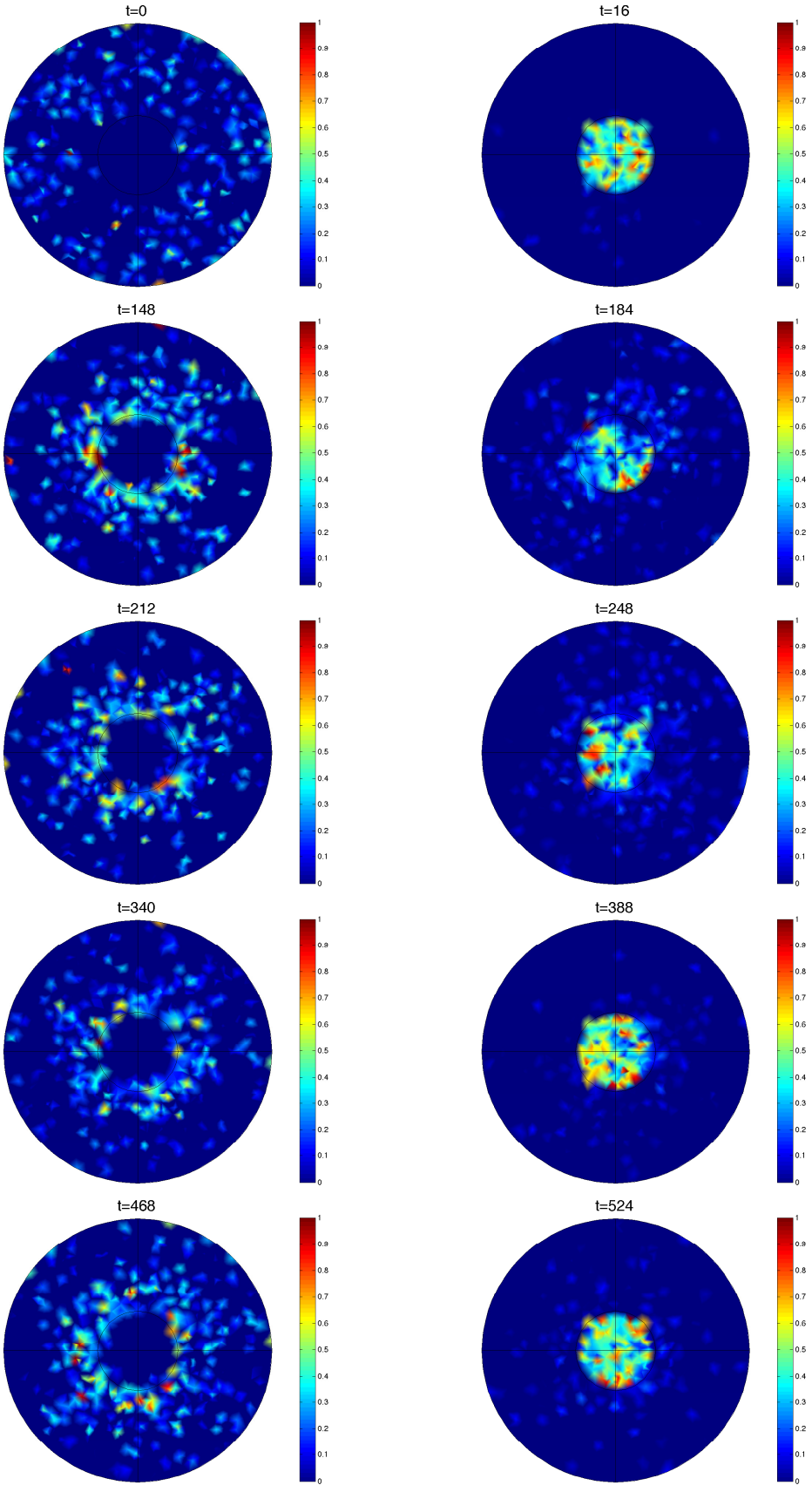

Figure 6.22: *Snapshots of NFκB. Gene sites clustered close to the nuclear membrane.* $D = 3 \times 10^{-11}$

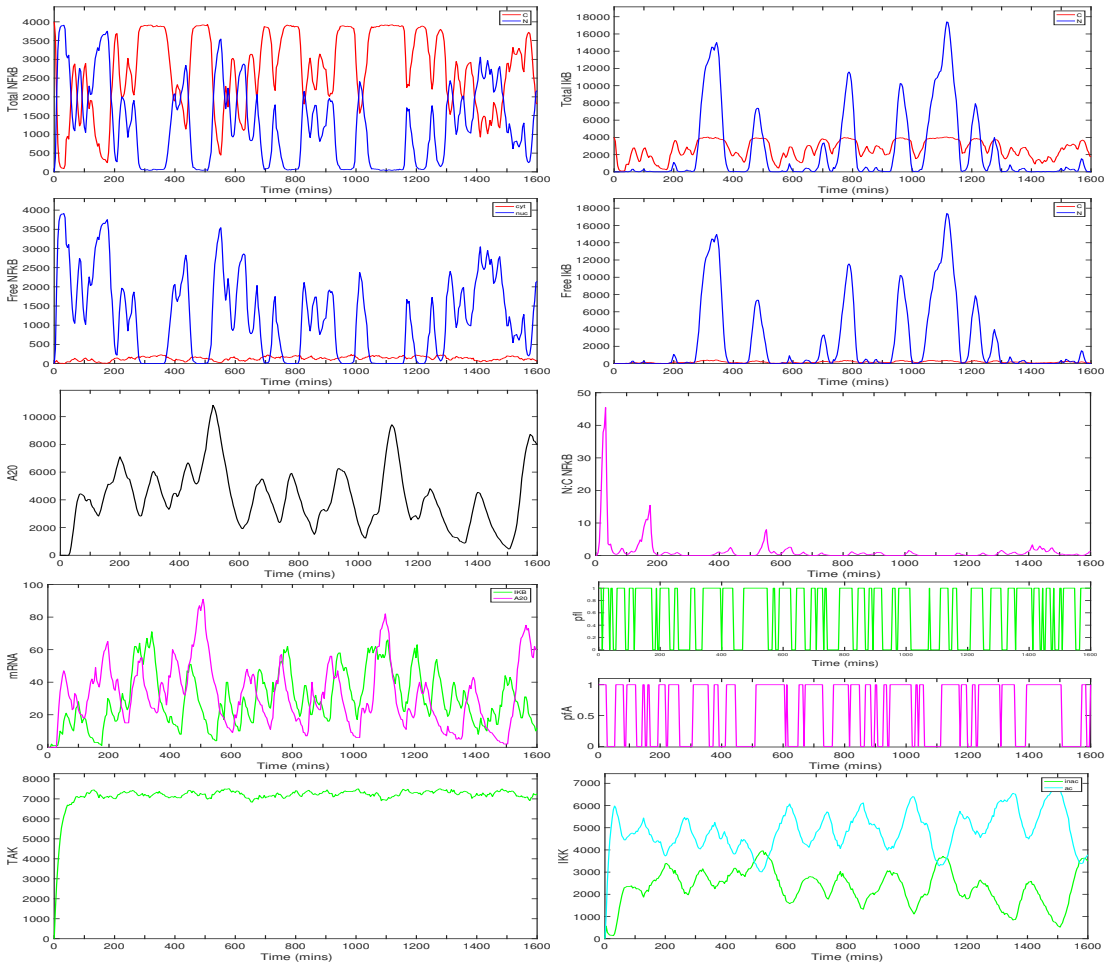

Figure 6.23: *Time series of species in the NFκB pathway. For each species, with the exception of NFκB and IκBα (in which we show nuclear and cytoplasmic numbers) the plots show the total copy number over the entire cell against time. Gene sites are clustered close to the nuclear membrane and* $D = 3 \times 10^{-11}$. *From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, Free IκB, A20, nuclear-cytoplasmic ration of NFκB, mRNA of IκB and A20, Free promoter species of IκB and A20, TAK, IKK (inactive and active).*

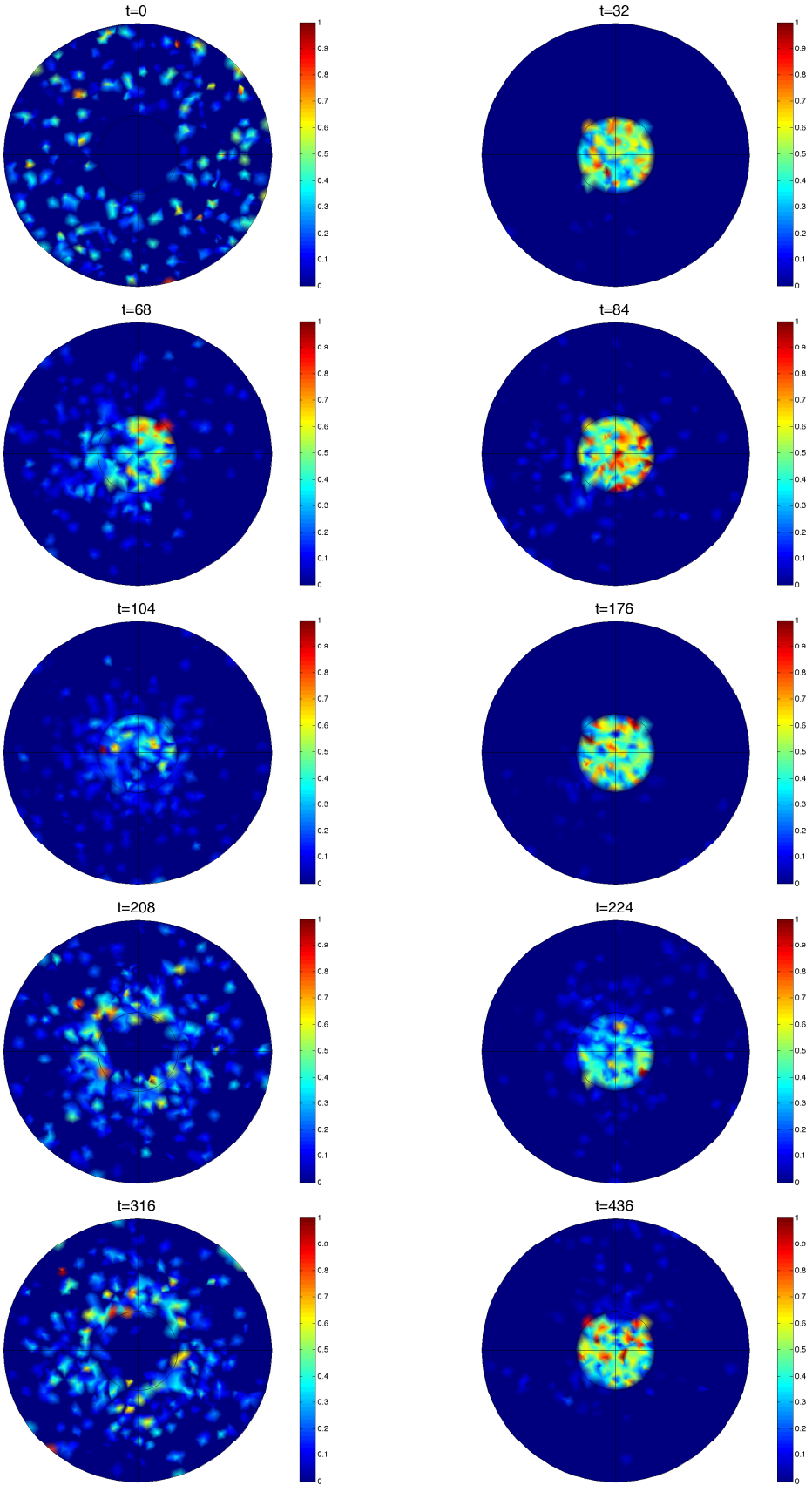

Figure 6.24: *Snapshots of NFκB. Gene sites clustered close to the nuclear membrane.* $D = 3 \times 10^{-11}$

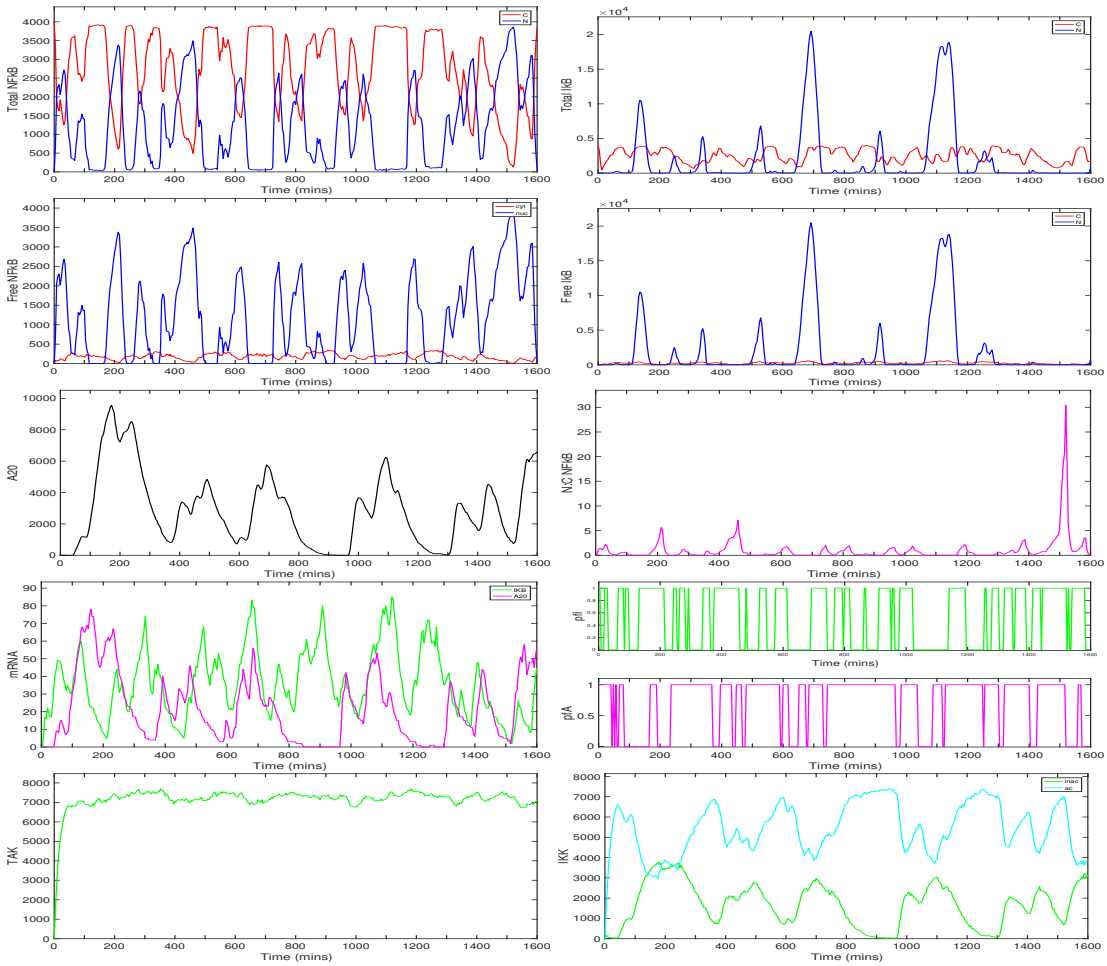

Figure 6.25: *Time series of species in the NFκB pathway. For each species, with the exception of NFκB and IκBα (in which we show nuclear and cytoplasmic numbers) the plots show the total copy number over the entire cell against time. Gene sites are clustered close to the nuclear membrane and* $D = 1 \times 10^{-10}$. *From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, Free IκB, A20, nuclear-cytoplasmic ration of NFκB, mRNA of IκB and A20, Free promoter species of IκB and A20, TAK, IKK (inactive and active).*

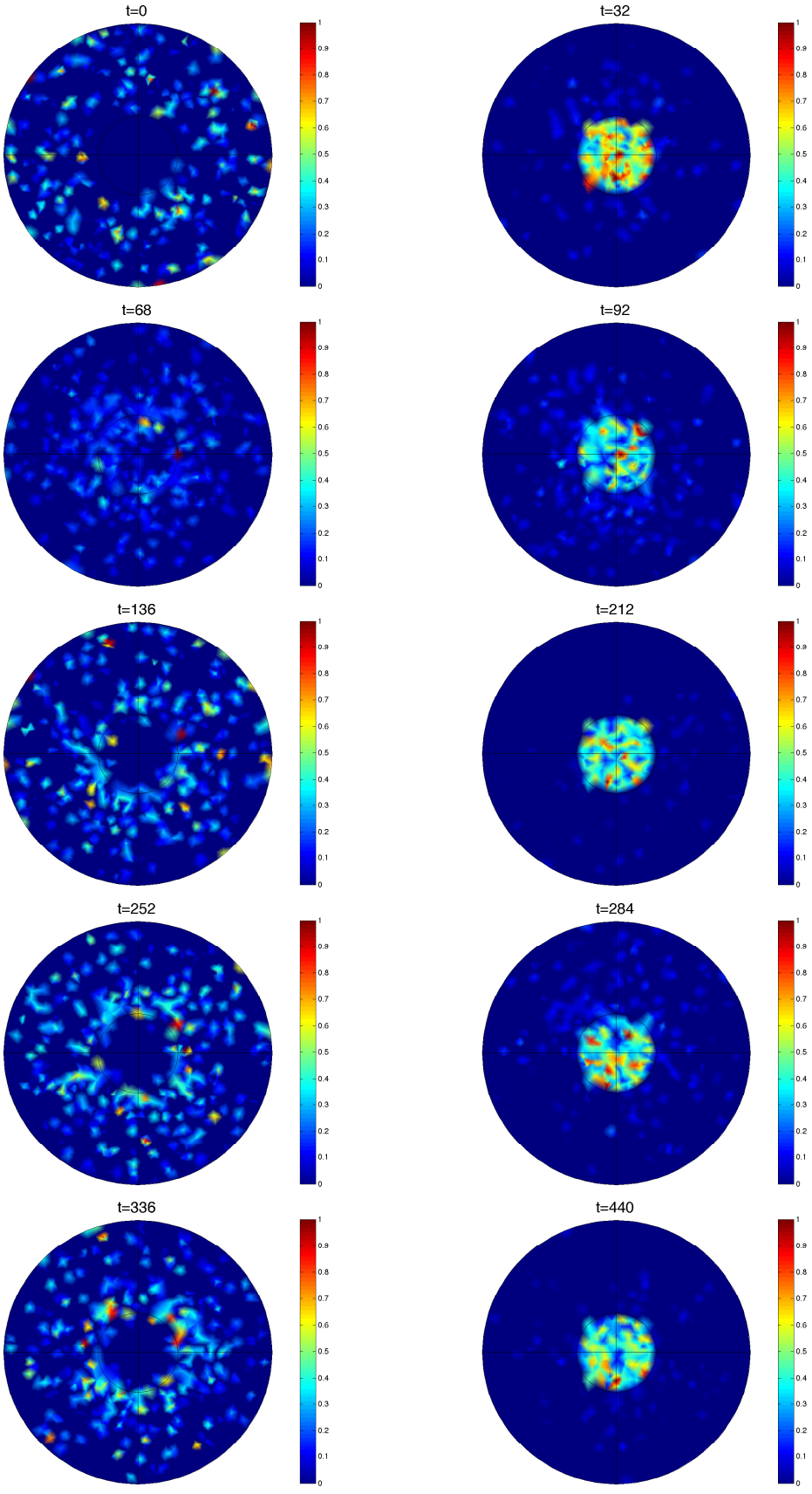

Figure 6.26: *Snapshots of NFκB. Gene sites clustered close to nuclear membrane* $D = 1 \times 10^{-10}$

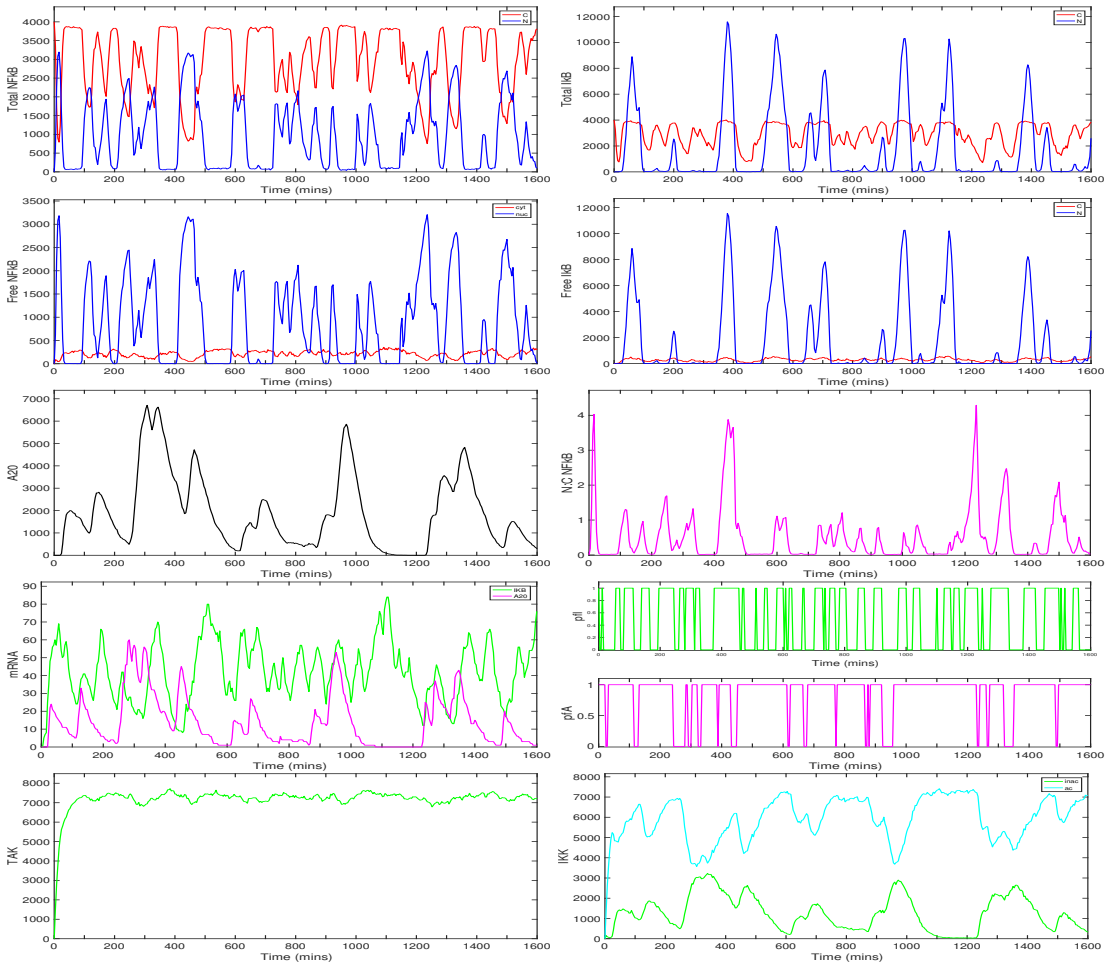

Figure 6.27: *Time series of species in the NFκB pathway. For each species, with the exception of NFκB and IκBα (in which we show nuclear and cytoplasmic numbers) the plots show the total copy number over the entire cell against time. Gene sites are clustered close to the nuclear membrane and* $D = 1 \times 10^{-10}$. *From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, Free IκB, A20, nuclear-cytoplasmic ration of NFκB, mRNA of IκB and A20, Free promoter species of IκB and A20, TAK, IKK (inactive and active).*

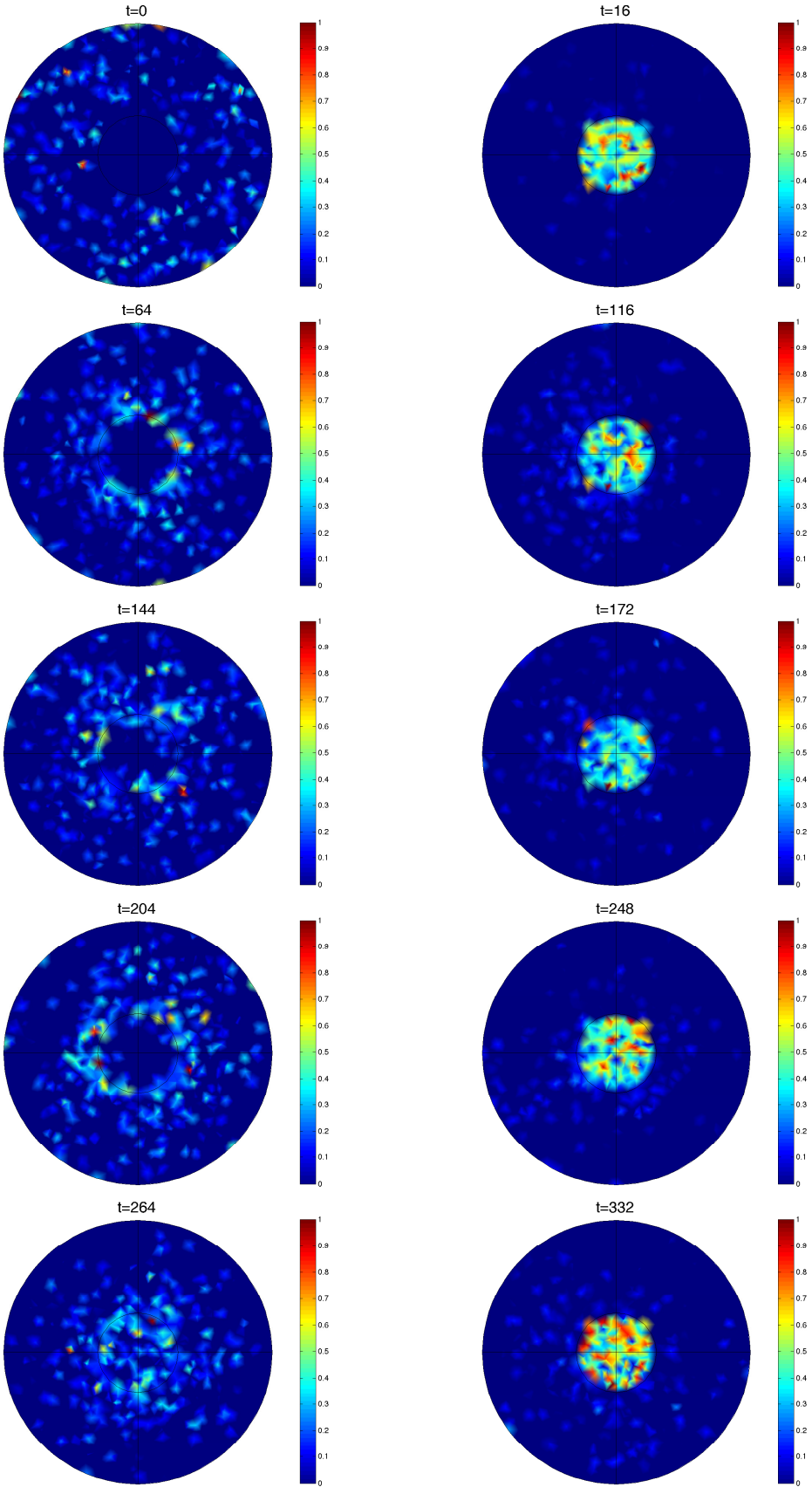

Figure 6.28: *Snapshots of NFκB. Gene sites clustered close to nuclear membrane* $D = 1 \times 10^{-10}$

For the slower diffusion regime, $D = 8 \times 10^{-12}$ an initial spike in nuclear NFκB was observed for most simulations, after this, NFκB went back to generally being held inactive in the cytoplasm by its inhibitor IκB, see figures 6.17, 6.18, 6.23 and 6.24. WAVOS detected average periods of 250 minutes in the nuclear NFκB species, which is a lot longer than experimentally observed. However, these nuclear NFκB oscillations were of a low amplitude, thus, only insignificant nuclear-cytoplasmic oscillations of NFκB were recorded. The closer proximity of the gene sites to the cytoplasm aids in NFκB reaching its target gene faster. This lead to a higher amount of fluctuations in NFκB-IκB and NFκB-A20 promoter binding and thus, the transcription of IκB and A20 mRNA. Due to the closer proximity of the gene sites to the cytoplasm, IκB and A20 mRNA can readily travel to the cytoplasm and IκB and A20 protein can be synthesised. Thus, there is a greater level in down-regulation of NFκB by its sequestration in the cytoplasm by its inhibitor IκB. A percentage of cells in experiments are observed to be inactive under TNF stimulation (Nelson et al. 2004). Our model here suggests that the position of the gene sites within those cells could be closer to the nucleus, thus, exhibiting inactivity.

However, as we increase the diffusion parameter from $D = 8 \times 10^{-12}$ to $D = 3 \times 10^{-11}$ and again to $D = 1 \times 10^{-10}$ we see the onset of oscillations. This difference can be explained by the higher amount of A20 that was seen to get synthesised. For this increase in the diffusion parameter we see an increase of 1000 of the peak numbers reached in A20 protein. Thus, we see more oscillatory behaviour in the IKK species, particularly more for the fast diffusion regime: see figures 6.21, 6.23, 6.25 and 6.27.

The mean period calculated for $D = 3 \times 10^{-11}$ was 178 minutes, ranging between 105 minutes and 266 minutes.

The mean period calculated for $D = 1 \times 10^{-10}$ was 131 minutes, ranging between 110 minutes and 166 minutes. For this diffusion regime, the copy numbers of nuclear

NFκB peak amplitudes also increased.

We then say with confidence, that diffusion drives oscillations in this model.

See figure 6.29 for a summary of the NFκB GRN. Figure 6.29 shows plots of mean period and mean copy number levels of nuclear NFκB against the diffusion parameter. For a diffusion parameter of $D = 1 \times 10^{-12}$ and gene sites clustered at the origin, we find NFκB to generally accumulate in the nucleus. Large mean periods were calculated but this was insignificant when it came to looking at the ratio of nuclear-cytoplasmic levels of NFκB. On the contrary, for a diffusion parameter of $D = 8 \times 10^{-12}$ and gene sites clustered close the nucleus membrane, we saw NFκB being held inactive in the cytoplasm. Generally we see that as diffusion is increased, there is an increase in the levels of nuclear NFκB oscillations.

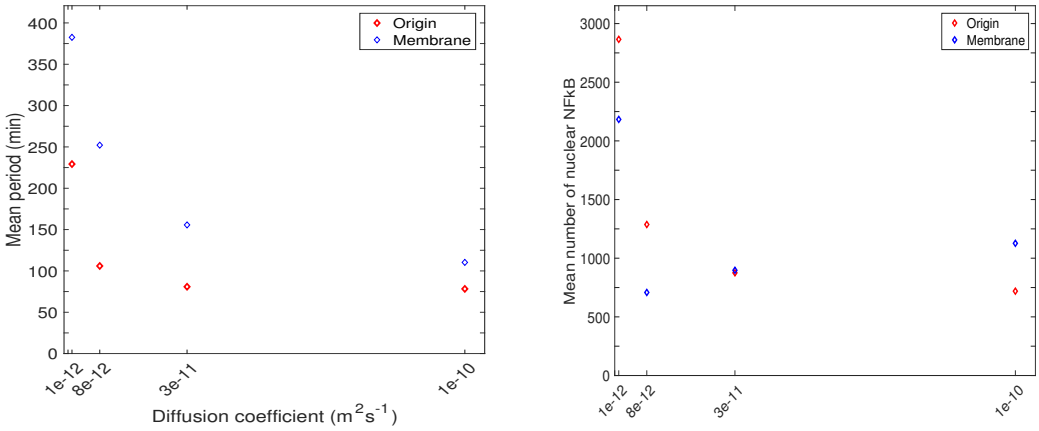

Figure 6.29: *Left: Average period recorded of nuclear NFκB oscillations over one hundred trajectories of the NFκB GRN against diffusion coefficient, D. Right: Average copy number of NFκB over one hundred simulations against diffusion coefficient. Both plots consider cases where gene sites are clustered together at the origin and the case where gene sites are clustered together close to the nuclear membrane.*

6.4 Crosstalk between NF-κB and HIF1α

HIF (Hypoxia Inducible Factor) is a family of heterodimeric transcription factors, that as the name suggests, play a major role in hypoxia, (low oxygen conditions). The pioneering investigations indeed found that a major function of HIF is to help achieve cellular homeostasis and survival during hypoxia.

The HIF dimers consist of an α subunit and a β subunit. The expression of HIFβ subunits are not affected by oxygen levels. However, HIF is usually only seen to be activated (stabilised) under hypoxic conditions. This is due to the stability of α subunits being heavily dependent on oxygen conditions. Under normoxic conditions, the HIFα subunits are normally seen to be broken down (Maynard et al. 2003). The hydroxylation of specific HIFα residues by PHDs (prolyl hydroxylases) stimulates interaction with the von Hippel-Lindau E3 Ligase, (VHL) resulting in the α subunit being targetted for disassembly.

However, recently HIF has been seen to be active in relatively well oxygenated regions of tumours and diseases such as rheumatoid arthritis and diabetes. These diseases are also common to NFκB.

NFκB and HIF share some target genes. Notably the pro-angiogenic cytokine $IL-8$ (Maxwell et al. 2007) and the inflammatory regulator, $TNF\alpha$ (Gorlach & Bonello 2008).

Due to these observations an intrinsic connection between the inflammatory and hypoxic responses was hypothesised and a relationship betweeen HIF and NFκB has now come to light.

6.4.1 HIF1α-NFκB crosstalk

Experiments have shown that if there is a complete block of the NFκB pathway during hypoxia, HIF1α is no longer activated and in fact HIF1α levels are diminished. Further investigation revealed the HIF1α promoter to be responsive to selective NFκB subunits, rendering NFκB a regulator of basal HIF1α mRNA transcription, (Bonello et al. 2007, Rius et al. 2008, van Uden et al. 2008). Additionally, despite the fact that HIFα subunits are normally degraded during normoxic conditions, it has recently come to pass that HIF1α can be stabilised (thus activated) during normoxia by NF-κB.

This regulation was demonstrated further via experiments that exposed cells to inflammatory cytokines such as TNFα, wherein, the subsequent NFκB activation induces HIF1α expression in an NFκB dependent manner (van Uden et al. 2008).

Moreover, Bandarra et al. (2015) unearthed the indirect suppression of NFκB by HIF1α via a mechanism that requires TAK-IKK activity. They found that in the absence of HIF1α, the levels of several NFκB targets, specifically A20, IκB, IL-8 and Cyld, (Renner & Schmitz 2009) either showed no significant reduction or were in fact elevated, due to them themselves being NFκB targets. This demonstrates the important role HIF plays in preventing excessive pro-inflammatory responses.

This confirms and elucidates futher the importance of controlling NFκB by the existence of many negative feedback loops that are in place within cells, HIF1α having recently been discovered to be one of them.

Thus, NFκB seems to be a vital player in the HIF GRN and vice versa.

6.4.2 Mathematical modelling of NF-κB and HIF1-α crosstalk

In the previous canonical NFκB GRN model we neglected the effects of HIF crosstalk. This was completely valid as termination of the NFκB response still occurs in the absense of HIF1α, due to the many other feedback mechanisms in place. Plus, it had been understood that HIF was degraded under normoxic conditions. We thus, now present a spatial stochastic model of the canonical NFκB GRN coupled with HIF1α. All the reactive and diffusive processes of the previous model still remain but now we include the addition of HIF1α effects. For instance, the indirect repression of NFκB by HIF1α is due to the interference of HIF1α with the TAK-IKK complex, decreasing the amount that IKK is activated and thus, decreasing the liberation of NFκB. We therefore, include a reaction where HIF1α forms a complex with the TAK-IKKi complex, preventing the activation of inactive IKK. Since, we now have NFκB as an inducer of HIF1α mRNA transcription, this completes the additional negative feedback loop.

In this section we will consider two cases of NFκB-HIF1α crosstalk. At first we will consider the case where the gene sites are clustered close to the nuclear membrane and second we will consider the case where the gene sites are clustered close to the nuclear membrane. We will stick to the one diffusion regime, where we set, $D = 8 \times 10^{-12}$.

For the first case, the position of the three gene sites are, $(x_{I\kappa B}, y_{I\kappa B}, z_{I\kappa B}) = (0.5\cos\pi//6\mu m, -0.5\sin$ for IκBα and $(x_{A20}, y_{A20}, z_{A20}) = (-0.5\cos\pi/6\mu m, -0.5\sin\pi/6\mu m, 0\mu m)$ for A20 and $(x_{HIF1\alpha}, y_{HIF1\alpha}, z_{HIF1\alpha}) = (0\mu m, 0.5\mu m, 0\mu m)$ for HIF1α. For the second case, the position of the three gene sites are translated by $2.35\mu m$ along the positive y-axis. The gene site positions for gene sites clustered close to the nucleus membrane are: $(x_{I\kappa B}, y_{I\kappa B}, z_{I\kappa B}) = (0.5\cos\pi//6\mu m, -0.5\sin\pi/6\mu m, 0\mu m)$ for IκBα and $(x_{A20}, y_{A20}, z_{A20}) =$ $(-0.5\cos\pi/6\mu m, -0.5\sin\pi/6\mu m, 0\mu m)$ for A20 and $(x_{HIF1\alpha}, y_{HIF1\alpha}, z_{HIF1\alpha}) = (0\mu m, 0.5\mu m, 0\mu m$ for HIF1α.

The first set of two simulations are for first set and the second set of two simulations are for the latter case. Again, each set comes in two: a time series of individual species copy number over the entire cell and a series of ten spatial snapshots of NFκB. Each snapshot is of a slice of the cellular domain, showing the x-y plane.

Table 6.5: *Cytoplasmic reactions*

Reaction	Description	Parameter value
Stimulus $\xrightarrow{\alpha}$ IKKi	Formation of inactive IKK	$\alpha = 0.0185 M min^{-1}$
Stimulus $\xrightarrow{\alpha}$ TAK	Formation of TAK	$\alpha = 0.0185 M min^{-1}$
IKKi + TAK $\xrightarrow{A_1}$ TAKIKKi	Formation of TAK and IKKi complex	$A_1 = 1 \times 10^{13} M^{-1} min^{-1}$
TAKIKKi $\xrightarrow{C_1}$ IKKa + TAK	Activation of IKK	$C_1 = 50 M^{-1} min^{-1}$
HIF1αtran $\xrightarrow{S_H}$ HIF1αtran + HIF1α	Translation of HIF1α protein	$S_H = 10 min^{-1}$
A20 + IKKa $\xrightarrow{A_3}$ IKKa-A20	Formation of IKKa-A20 complex	$A_3 = 1 \times 10^{11} M^{-1} min^{-1}$
IKKaA20 $\xrightarrow{C_3}$ IKKi + A20	Deactivation of IKKa via interaction with A20	$C_3 = 50 min^{-1}$
IKKi $\xrightarrow{d_k} \emptyset$	Degradation of IKKa	$d_k = 0.06 min^{-1}$
$IKKa \xrightarrow{D_k}$ IKKi	Spontaneous deactivation of IKKa	$D_k = 0.03 min^{-1}$
HIF1α + TAKIKKi $\xrightarrow{A_4}$ TAKIKKiHIF	Association/disassociation of TAK-HIF complex	$A_4 = 1 \times 10^{1}1 M^{-1} min^{-1}$
TAKIKKiHIF $\xrightarrow{C_4}$ TAK + IKKi + HIF1α	Disassociation of TAKIKKiHIF complex	$d_k = 50 min^{-1}$
HIF1α $\underset{H\sigma_{off}}{\overset{H\sigma_{on}}{\rightleftharpoons}}$ HIF1αmic	HIF1αB binding/unbinding to microtubule	$H\sigma_{on} = 1 \times 10^{9} min^{-1}$ $H\sigma_{off} = 10 min^{-1}$
$H\sigma_i \xrightarrow{v} H\sigma_j$	Radially directed active transport of HIF1α between connected voxels	$v = 3 \times 10^{-5} m min^{-1}$
$S_{ij} \xrightarrow{d_{jik}} S_{ik}$	Molecular diffusion	$D = 5 \times 10^{-11} m^2 min^{-1}$

Table 6.6: *HIF1α gene site reactions*

Reaction	Description	Parameter value
$Hp_f \xrightarrow{\alpha_H} Hp_f$ + HIF1αtran	Basal transcription of HIF1α mRNA	$\alpha_H = 2.25 min^{-1}$
$Hp_o \xrightarrow{\alpha_H/\gamma} Hp_o + HIF1\alpha$tran	Repression of HIF1α mRNA transcription	$\alpha_H = 2.25\,\text{min}^{-1}, \gamma = 100$
$Hp_f + NF\kappa B \underset{b_2}{\overset{b_1}{\rightleftharpoons}} Hp_o$	NFκB binding/unbinding to HIF1α promoter	$b_1 = 3.35 \times 10^6 min^{-1}$ $b_2 = 0.07 min^{-1}$

Table 6.7: *Global reactions*

Global Reactions	Description	Parameter value
HIFtran $\xrightarrow{d_{mH}} \emptyset$	Degradation of HIFtran	$d_m = 0.025 min^{-1}$
HIF $\xrightarrow{d_{pH}} \emptyset$	Degradation of HIF	$d_p = 0.07 min^{-1}$

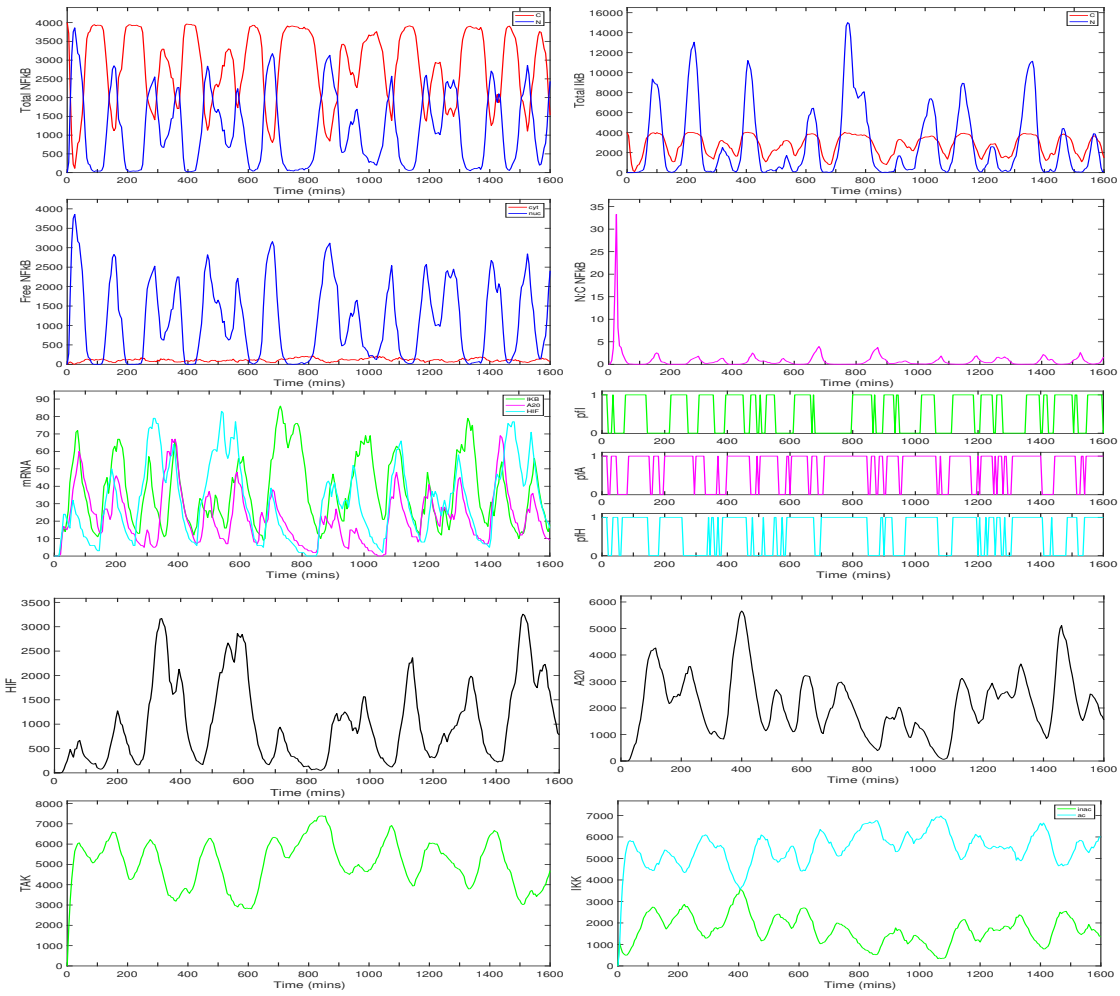

Figure 6.30: *Time series of species in the NFκB pathway with HIF1α pathway. For each species, with the exception of NFκB and IκBα (in which we show nuclear and cytoplasmic numbers) the plots show the total copy number over the entire cell against time. Gene sites are clustered close to the centre of the nucleus and* $D = 8 \times 10^{-12}$*. From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, nuclear-cytoplasmic ration of NFκB, mRNA of IκB, A20 and HIF1α, Free promoter species of IκB, A20 and HIF1α, HIF1α, A20, TAK and IKK (inactive and active).*

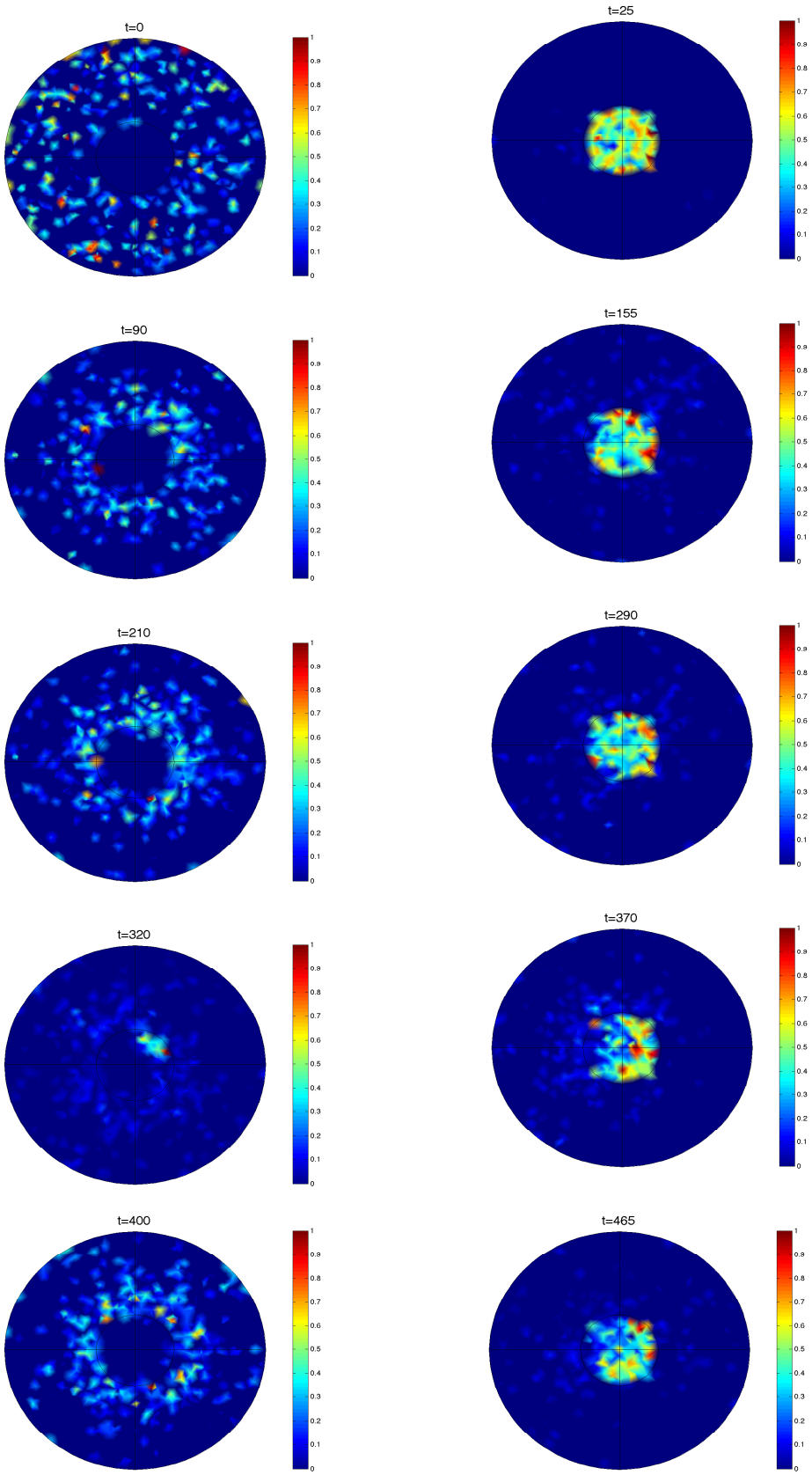

Figure 6.31: *Snapshots of NFκB: Crosstalk with HIF1α. Gene sites are clustered close to the origin.* $D = 8 \times 10^{-12}$

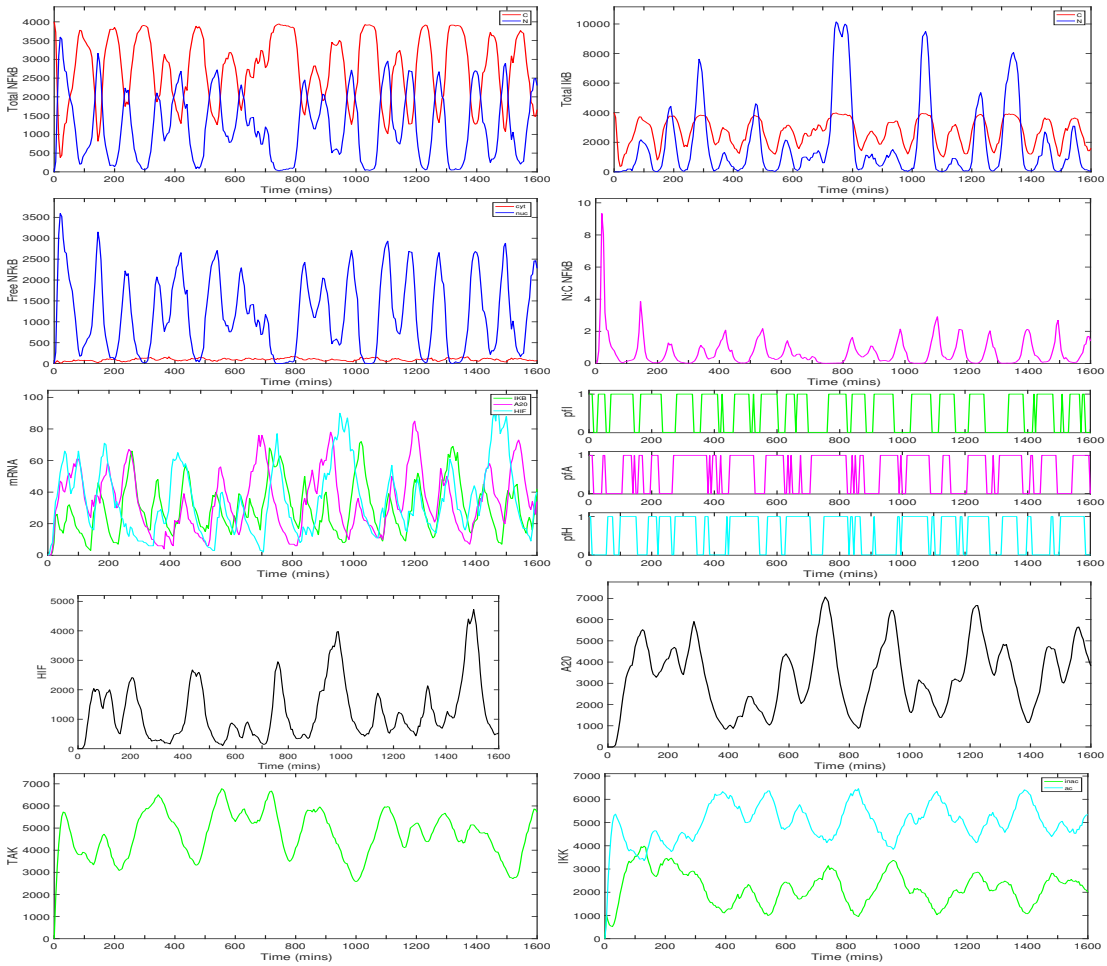

Figure 6.32: *Time series of species in the NFκB pathway with HIF1α pathway. For each species, with the exception of NFκB and IκBα (in which we show nuclear and cytoplasmic numbers) the plots show the total copy number over the entire cell against time. Gene sites are clustered close to the centre of the nucleus and* $D = 8 \times 10^{-12}$. *From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, nuclear-cytoplasmic ration of NFκB, mRNA of IκB, A20 and HIF1α, Free promoter species of IκB, A20 and HIF1α, HIF1α, A20, TAK and IKK (inactive and active).*

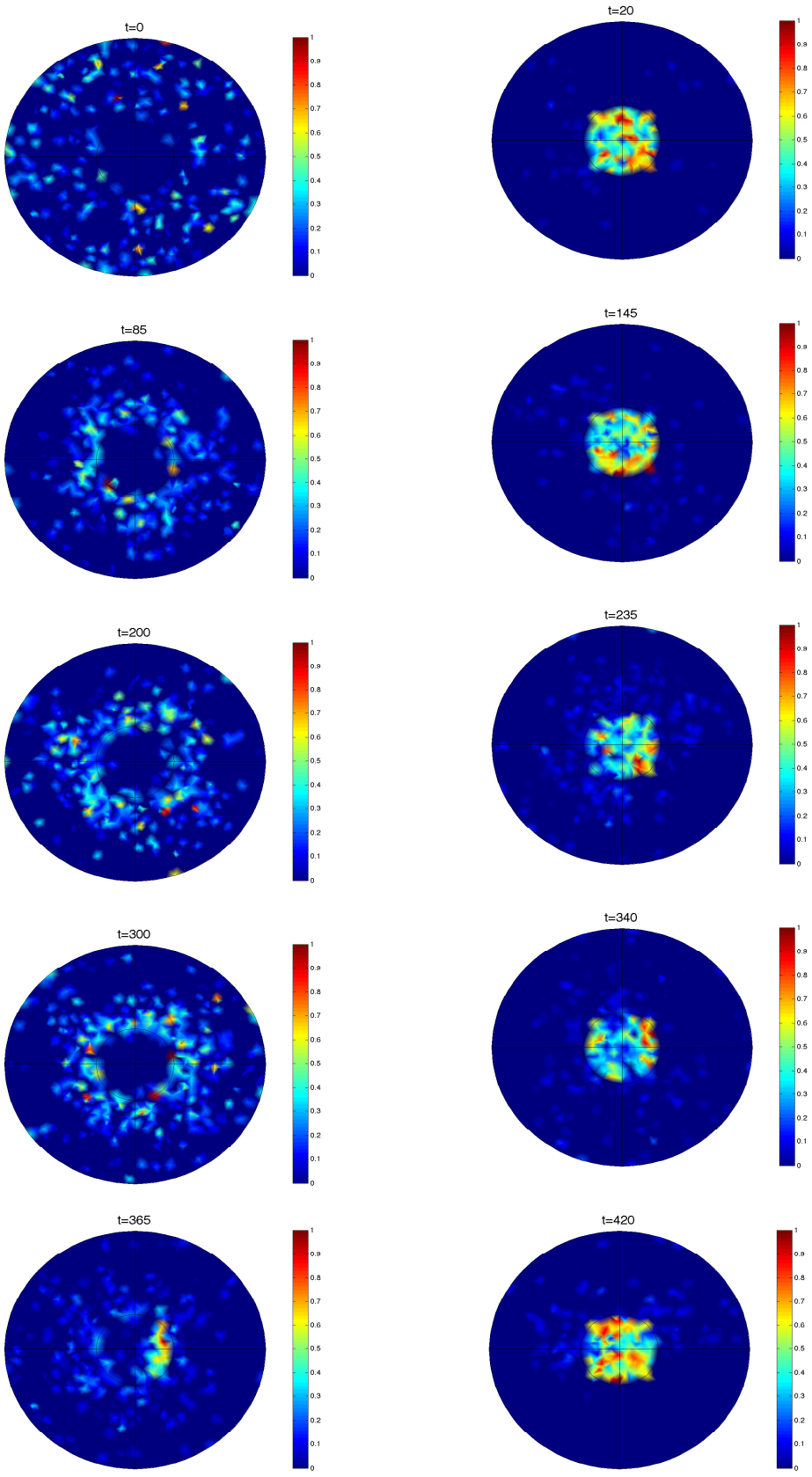

Figure 6.33: *Snapshots of NFκB: Crosstalk with HIF1α. Gene sites are clustered close to the origin.* $D = 8 \times 10^{-12}$

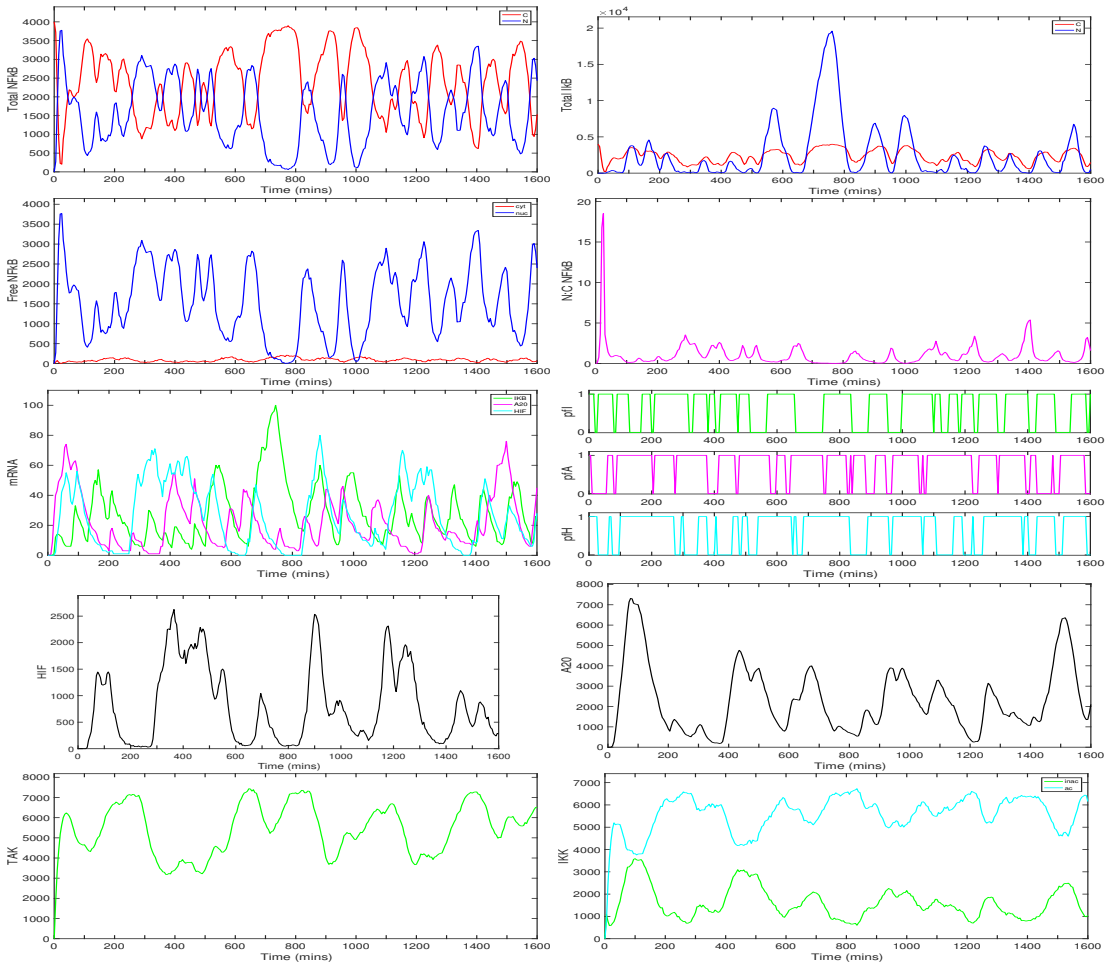

Figure 6.34: *Time series of species in the NFκB pathway with HIF1α pathway. For each species, with the exception of NFκB and IκBα (in which we show nuclear and cytoplasmic numbers) the plots show the total copy number over the entire cell against time. Gene sites are clustered close to the nuclear membrane and* $D = 8 \times 10^{-12}$*. From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, nuclear-cytoplasmic ration of NFκB, mRNA of IκB, A20 and HIF1α, Free promoter species of IκB, A20 and HIF1α, HIF1α, A20, TAK and IKK (inactive and active).*

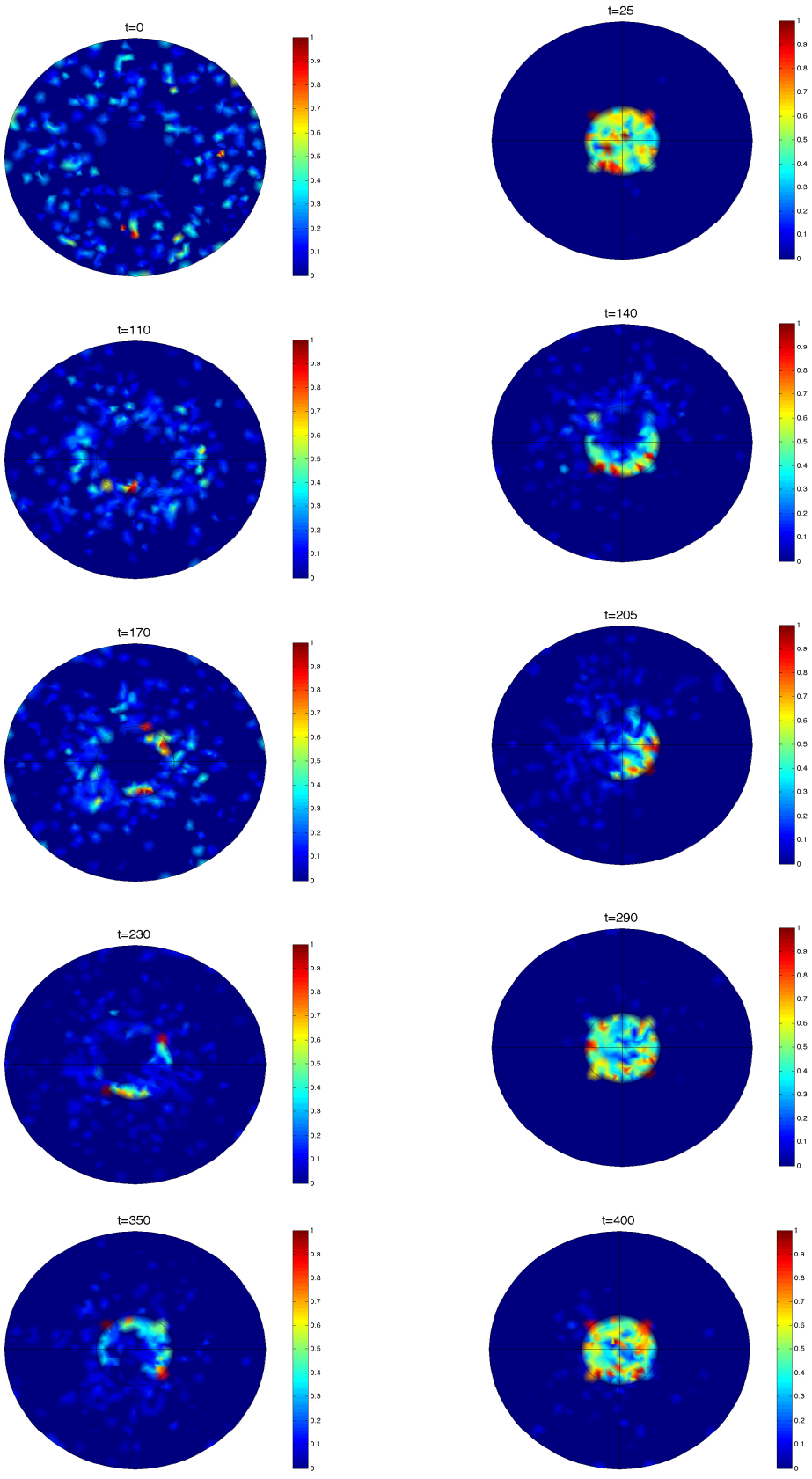

Figure 6.35: *Snapshots of NFκB: Crosstalk with HIF1α. Gene sites are clustered close to the nuclear membrane.* $D = 8 \times 10^{-12}$

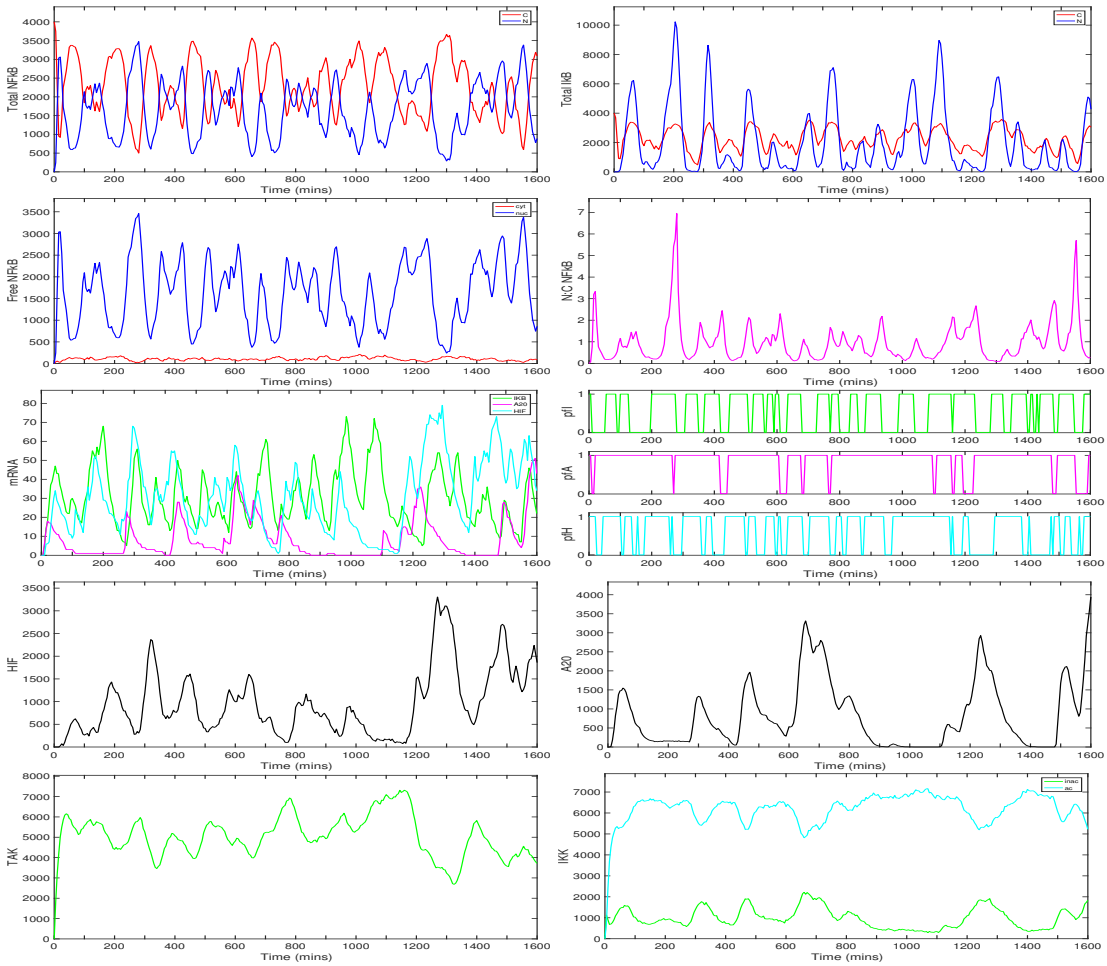

Figure 6.36: *Time series of species in the NFκB pathway with HIF1α pathway. For each species, with the exception of NFκB and IκBα (in which we show nuclear and cytoplasmic numbers) the plots show the total copy number over the entire cell against time. Gene sites are clustered close to the nuclear membrane and* $D = 8 \times 10^{-12}$*. From top left down to bottom right: Total NFκB, Total IκB, Free NFκB, nuclear-cytoplasmic ration of NFκB, mRNA of IκB, A20 and HIF1α, Free promoter species of IκB, A20 and HIF1α, HIF1α, A20, TAK and IKK (inactive and active).*

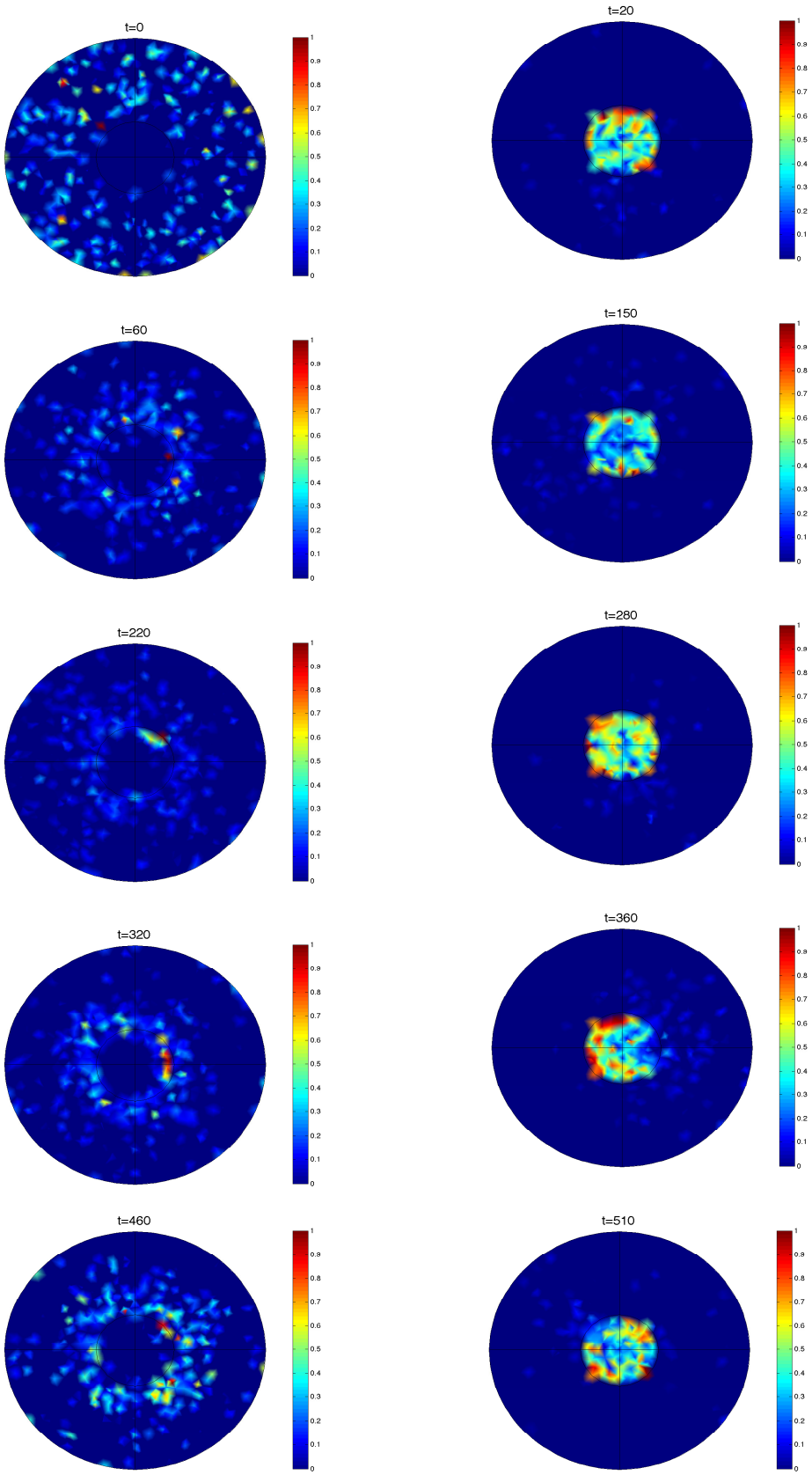

Figure 6.37: *Snapshots of NFκB: Crosstalk with HIF1α. Gene sites are clustered close to the nuclear membrane.* $D = 8 \times 10^{-12}$

With the addition of HIF1α crosstalk, the period of nuclear-cytoplasmic oscillations of NFκB is ever so slightly increased but not significantly, (for the case where gene sites are clustered close to the origin). However, an interesting result, is in the comparison between the NFκB GRN with gene sites clustered close to the nuclear membrane, with and without HIF. See figures 6.21, 6.22, 6.23, 6.24, for the NFκB GRN dynamics in the absence of HIF1α and figure 6.34, 6.35, 6.36 and 6.37 with the inclusion of HIF. The fact that the gene sites were closer to the nuclear membrane again lead to the greater number of protein being synthesised. However, the explanation for the onset of oscillations at this position under this diffusion regime was the inclusion of HIF1α. The extra negative feedback to the system, induced the oscillatory nature in this case.

We speculate that the presence of heterogeneity throughout a cell population, in terms of the observation of NFκB activation, could be due to the positioning of gene sites within the cell and/or whether or not HIF1α is active in a cell.

The restriction our model put on crossing the nuclear membrane was of the utmost importance in creating the higher values of nuclear-cytoplasmic ratio. Also, if active transport was turned of we noticed how similar results could still be recorded but only for an increase in the diffusion parameter.

One remark, we can make is that the comparison made between simulation and experiment was vastly done by visual observation. By computing the power spectrum for both experiments and simulations, a truly quantitative comparison can be carried out. This will be left for future work.

Spatial stochastic modelling was essential in capturing the variability in key aspects of GRN dynamics.

Chapter 7

Discussion and future work

7.1 Discussion

In this thesis we have focussed only on the behaviour of intracellular gene regulation in isolation within a single cell. Of course in many *in vitro* experiments and *in vivo*, cells exist in communities and it is important that they communicate with and signal to each other. For example, as is the case for the developmental process of somitogenesis, intracellular signalling is coordinated in an intercellular manner. In particular the oscillatory expression of certain proteins within cells can be synchronous (see, for example, Lewis 2003, Terry & Chaplain 2011). Understanding how cell-cell interactions affect gene regulation and the dynamics of a group of cells or a tissue could form the basis of future investigation.

Overall the results of this thesis confirm the importance of modelling transcription factor systems, where negative feedback loops are involved (both actual and synthetic), using explicit spatial models. Given the current level of interest in synthetic biology

and the technological tools available to synthetic biologists, the findings in this thesis indicate that experimentalists should take molecular movement into account when trying to design such systems.

Most of the previous work in this area has adopted a delay differential equation approach where a discrete time delay is included in a system of ODEs. As well as transcription and translation, the delay is taken to account for molecular movement without explicitly incorporating this mechanism into the model. By explicitly incorporating spatial terms into our model, we are able to say something more focused about the importance of molecular movement and the effect of molecular transport time between nucleus and cytoplasm. Recent work on spatial models and the results of this thesis show that the incorporation of spatial effects into models of GRNs allows us to reproduce the known oscillatory dynamics. While delay equations also capture the overall oscillatory dynamics of GRNs, considering spatial models which incorporate intracellular molecular movement directly, will allow connections to be made with experimental data arising from single cell experiments. Increasingly biologists are developing techniques to tag and monitor the movement of molecules in single cells. Developing appropriate mathematical models that have the ability to analyse the spatial data that arises from such single-cell experiments is also, therefore, important, which is where spatial models such as those presented in this thesis can bring new insights to the problem. In addition to describing the overall mRNA and protein concentrations over time the computational results of our model may be compared with single cell experiments where proteins are tagged with a fluorescent label, (e.g. Ashall et al. 2009, Lahav et al. 2004, Nelson et al. 2004), although this is beyond the remit of this thesis. Populations of cells are heterogenous in nature, with differences at both the genetic and the phenotypic level. In order to continue to study intracellular dynamics (and potentially the subsequent cell-cell dynamics) it is important to have

mathematical models which can account explicitly for the phenotypic variation between individual cells. Spatial models permit a realistic modelling of individual cells where it is required to analyse the aspects of phenotypic variation which arise from differences in intracellular structure – such as different positioning of gene sites, variations in diffusion coefficients between molecules, spatially-dependent diffusion due to intracellular structural heterogeneity, or explicit transport of molecules across the nuclear membrane. As the imaging techniques themselves are further developed and refined it is also important to continue to develop spatial mathematical models.

Bibliography

Agrawal, S., Archer, C. & Schaffer, D. V. (2009), 'Computational models of the notch network elucidate mechanisms of context-dependent signaling', *PLoS Computational Biology* **5**, 1–14.

Alon, U. (2007), 'Network motifs: theory and experimental approaches', *Nature* **8**, 450–461.

Arenzana-Seisdedos, F., Turpin, P., Rodriguez, M., Thomas, D., Hay, R. T., Virelizier, J. & Dargemont, C. (1997), 'Nuclear-localisation of IκBα promotes active transport of NF$-\kappa$B from the nucleus to the cytoplasm', *Journal of Cell Science* **110**, 369–378.

Ashall, L., Horton, C. A., Nelson, D. E., Paszek, P., Harper, C. V., Sillitoe, K., Ryan, S., Spiller, D. G., Unitt, J. F., Broomhead, D. S., Kell, D. B., Rand, D. A., See, V. & White, M. R. H. (2009), 'Pulsatile stimulation determines timing and specificity of NF-κB-dependent transcription', *Science* **324**, 242–246.

Bagnall, J., Leedale, J., Taylor, S. E., Spiller, D. G., White, M. R. H., Sharkey, K. J., Bearon, R. N. & See, V. (2013), 'Tight control of hypoxia-inducible factor-α transient dynamics is essential for cell survival in hypoxia', *The journal of biological chemistry* **289**, 5549–5564.

Balagadde, F. K., Song, H., Ozaki, J., Collins, C. H., Barnet, M., Arnold, F. H., Quake,

S. R. & You, L. (2008), 'A synthetic eschericia coli predator-prey ecosystem.', *Mol. Sys. Biol.* **4**(187).

Bandarra, D., Biddlestone, J., Mudie, S., Muller, H. A. J. & Rocha, S. (2015), 'Hif-1α restricts nf-κb-dependent gene expresssion to control innate immunity signals', *Disease Models and Mechanisms* **8**, 169–181.

Barrio, M., Burrage, K., Leier, A. & Tain, T. (2006), 'Oscillatory regulation of hes1: discrete stochastic delay modelling and simulation', *PLoS Computational Biology* **2**, 1017–1030.

Becskei, A., Kaufmann, B. & van Oudenaarden, A. (2005), 'Contribution of low molecule number and the chromosomal positioning to stochastic gene expression', *Nature Genetics* **37**, 937–944.

Becskei, A. & Serrano, L. (2000), 'Engineering stability in gene networks by autoregulation', *Nature* **405**, 590–593.

Behar, M. & Hoffmann, A. (2010), 'Understanding the temporal codes of intra-cellular singals', *Current opinions in genetics and development* **20**, 684–693.

Behar, M. & Hoffmann, A. (2013), 'Tunable signal processing through a kinase control cycle: the ikk signaling node', *Biophysical Journal* **105**, 231–241.

Bell-Pedersen, D., Cassone, V. M., Earnest, D. J. & Zoran, M. J. (2005), 'Circadian rhythms from multiple oscillators: lessons from diverse organisms', *Nat. Rev. Gen.* **6**, 544–556.

Bennet, M. R., Volfson, D., Tsimring, L. & Hasty, J. (2007), 'Transient dynamics of genetic regulatory networks', *Biophys. J.* **92**, 3501–3512.

Bernard, S., Cajavec, B., Pujo-Menjouet, L., Mackey, M. C. & Herzel, H. (2006), '2006', *Philos. Trans. A. Math. Phys. Eng. Sci.* **15**, 1155–1170.

Berridge, M. J., Bootman, M. D. & Roderick, H. L. (2003), 'Calcium signalling: Dynamics, homeostasis and remodelling', *Nature reviews: Molecular Cell Biology* **4**, 517–529.

Bonello, S., Zahringer, C., BelAiba, R. S., Djordjevic, T., Hess, J., Michiels, C., Kietzmann, T. & Gorlach, A. (2007), 'Reactive oxygen species activate the hif1α promoter via a functional nfκb site', *Arterioscler Thromb Vasc Biol* **27**, 755–761.

Bratsum, D., Volfson, D., Tsimring, L. S. & Hasty, J. (2005), 'Delay-induced stochastic oscillations in gene regulation', *Proceedings of the National Academy of Sciences* **102**, 14593–14598.

Busenberg, B. & Mahaffy, J. M. (1985), 'Interaction of spatial diffusion and delays in models of genetic control by repression.', *J. Math. Biol.* **22**, 313–333.

Carr, A. J. & Whitmore, D. (2005), 'Imaging of single light-responsive clock cells reveals fluctuating free-running periods', *Nature Cellular Biology* **7**, 319–321.

Chaplain, M. A. J., Ptashnyk, M. & Sturrock, M. (2014), 'Hopf bifurcation in a gene regulatory network model: Molecular movement causes oscillations', *Mathematical Models and Methods in Applied Science* **8**(30).

Chen, Y. Y., Galloway, K. E. & Smolke, C. D. (2012), 'Synthetic biology: advancing biological frontiers by building synthetic systems.', *Genome Biol.* **13**(240).

Cole, C. N. & Scarcelli, J. J. (2006), 'Transport of messenger rna from the nucleus to the cytoplasm', *Current Opinion in Cell Biology* **18**, 299–306.

Dequeant, M., Glynn, E., Gaudenz, K., Wahl, M., Chen, J., Mushegian, A. & Pourquie, O. (2006), 'A complex oscillatory network of signalling genes underlies the mouse segmentation clock', *Science* **314**, 1595–1598.

Dolmetsch, R. E., Lewis, R. S., Goodnow, C. C. & Hardy, J. I. (1997), 'Differential activation of transcription factors induced by Ca^{2+} response amplitude and duration', *Nature* **386**, 855–859.

Dolmetsch, R. E., Xu, K. & Lewis, R. S. (1998), 'Calcium oscillations increase the efficiency and specificity of gene expression', *Nature* **392**, 933–936.

Drawert, B., Engblom, S. & Hellander, A. (2012), 'Urdme: a modular framework for stochastic simulation of reaction-transport processes in complex geometries', *BMC Syst. Biol.* **6**(76).

Eldar, A. & Elowitz, M. B. (2010), 'Functional roles for noise in genetic circuits', *Nature* **467**, 167–173.

Elf, J. & Ehrenberg, M. (2004), 'Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases', *Syst. Biol.* **1**(2), 230–236.

Elowitz, M. B. & Leibler, S. (2000), 'A synthetic oscillatory network of transcriptional regulators', *Nature* **403**, 335–338.

Engblom, S., Ferm, L., Hellander, A. & Lotstedt, P. (2009), 'Simulation of stochastic reaction-diffusion processes on unstructured meshes', *SIAM J. Sci. Comput.* **31**(3), 1774–1797.

Fakhouri, W. D., Ay, A., Sayal, R., Dresch, J., Dayringer, E. & Arnosti, D. N. (2010), 'Deciphering a transcriptional regulatory code: nodeling short-range repression in the drosophila embryo', *Mol. Sys. Biol.* **6**(341).

Fange, D. & Elf, J. (2006), 'Noise-induced min phenotypes in e. coli.', *PLoS Comput. Biol.* **2**(e80).

Feller, W. (1950), *Introducation to probability theory and its applications*, J. Wiley, New York.

Fisher, R. A. (1958), *The genetical theory of natural selection*, Dover Publics., New York.

Forsythe, J. A., Jiang, B., Iyer, N. V., Agani, F., Leung, S. W., Koos, R. D. & Semenza, G. L. (1996), 'Activation of vascular endothelial growth factor gene transcription by hypoxia-inducible factor 1', *Molecular and Cellular Biology* pp. 4604–4613.

Gardner, T. S., Cantor, C. R. & Collins, J. J. (2000), 'Construction of a genetic toggle switch in escherichia coli', *Nature* **403**, 339–342.

Gibson, M. A. & Bruck, J. (2000), 'Efficient exact stochastic simulation of chemical systems with many species and many chennels', *J. Phys. Chem.* **104**, 1876–1889.

Gillespie, D. T. (1977), 'Exact stochastic simulation of coupled chemical reactions.', *J. Phys. Chem.* **81**, 2340–2361.

Glass, L. & Kauffman, S. A. (1974), 'Co-operative components, spatial localisation and oscillatory cellular dynamics', *J. Theor. Biol* **34**, 219–237.

Goodwin, B. C. (1965), 'Oscillatory behaviour in enzymatic control processes', *Adv. Enzyme Regul.* **3**, 425–428.

Gorlach, A. & Bonello, S. (2008), 'The cross-talk between nf-κb and hif-1: further evidence for a significant liaison', *J. Biochem* **412**, 17–19.

Griffith, J. S. (1968), 'Mathematics of cellular control processes. i. negative feedack to one gene.', *J. Theor. Biol.* **20**, 202–208.

Guck, J., Schinkinger, S., Lincoln, B., Wottawah, F., Ebert, S., Romeyke, M., Lenz, D., Erickson, H. M., Ananthakrishnan, R., Mitchell, D., Kas, J., Ulvick, S. & Bilby, C. (2005), 'Optical deformability as an inherent cell marker for testing malignant transformation and metastatic competence', *Biophys J.* **88**, 3689–3698.

Haggie, P. M., Stanton, B. A. & Verkman, A. S. (2004), 'Increased diffusional mobility of cftr at the plasma membrane after deletion of its c-terminal pdz binding motif', *J. Biol. Chem* **279**, 5494–5500.

Harang, R. & Petzold, L. R. (2012), 'Wavos: a matlab toolkit for wavelet analysis and visualization of oscillatory systems', *BMC Res. Notes* **5**(163).

Hayden, M. S. & Ghosh, S. (2012), 'NFκB, the first quarter-century: remarkable progress and outstanding questions', *Genes and Development* **26**, 203–234.

Higham, D. J. (2007), 'Modelling and simulating chemical reactions', *Society for Industrial and Applied Mathematics Review* **50**, 347–368.

Hilbert, L., Albrecht, D. & Mackey, M. C. (2011), 'Small delay, big waves: a minimal delayed negative feedback model captures eschericia coli single cell sos kinetics.', *Mol. Biosyst.* **7**(9), 2599–2607.

Hirata, H., Yoshiura, S., Ohtsuka, T., Bessho, Y., Harada, T., Yoshikawa, K. & Kageyama, R. (2002), 'Oscillatory expression of the bhlh factor hes1 regulated by a negative feedback loop', *Science* **298**, 840–843.

Hirota, K. & Semenza, G. L. (2006), 'Regulation of angiogensis by hypoxia-inducible factor 1', *Critical Reviews in Oncology/Hematology* **59**, 15–26.

Hoffmann, A., Levchenko, A., Scott, M. L. & Baltimore, D. (2002), 'The IκB – NFκB signalling module: temporal control and selective gene activation', *Science* **298**, 1241–1245.

Howard, H. & Rutenberg, A. D. (2003), 'Pattern formation inside bacteria: fluctuations due to the low copy number of proteins.', *Phys. Rev. Lett.* **90**(128102).

Huang, L., Yuan, Z., Liu, P. & Zhou, T. (2014), 'Feedback-induced counterintuitive correlations of gene expression noise with bursting kinetics', *Phys. Rev. E* **90**.

Imayoshi, I. & Kageyama, R. (2014), 'bhlh factors in self-renewal, multipotency, and fate choice of neural progenitor cells', *Neuron* **82**(1), 9–23.

Jaeger, J., Surkova, S., Blagov, M., Janssens, H., Kosman, D., Kozlov, K. N., Manu, Myasnikova, E., Vanario-Alonso, C. E., Samsonova, M., Sharp, D. H. & Reinitz, J. (2004), 'Dynamic control of positional information in the early drosophila blastoderm', *Nature* **430**, 368–371.

Jaenisch, R. & Bird, A. (2003), 'Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals', *Nature Genetics Supplement* **33**, 245–254.

Jensen, M. H., Sneppen, K. & Tiana, G. (2003), 'Sustained oscillations and time delays in gene expression of protein hes1', *FEBS Letters* **541**, 176–177.

Johansson, T., Lejonklou, M. H., Ekeblad, S., Stalberg, P. & Skogseid, B. (2008), 'Lack of nuclear expression of hairy and enhancer of split-1 (hes1) in pancreatic endocrine tumors', *Horm. Matab. Res.* **40**, 354–359.

Kageyama, R., Shimojo, H. & Imayoshi, I. (2015), 'Dynamic expression and roles of hes factors in neural development', *Cell Tissue Research* **359**, 125–133.

Kemler, I. & Fontana, A. (1999), 'Roles of iκbα and iκbβ in the biphasic nuclear translocation of nf-κb in tnfα-stimulated astrocytes and in neuroblastoma cells', *GLIA* **26**, 212–220.

Kepler, T. B. & Elston, T. C. (2001), 'Stochasticity in transcriptional regulation: origins consequences, and mathematical representations', *Biophys. J.* **81**, 3116.

Khokhar, J. Y., Ferguson, C. S., Zhu, A. Z. X. & Tyndale, R. F. (2010), 'Pharmocogenetics of drug dependence: Role of gene variation in susceptibility and treatment', *Annual Review Pharmacological Toxicology* **50**, 39–60.

Kim, I. S., Kim, D. H., Han, S. M., Chin, M. U., Nam, H. J., Cho, H. P., Choi, S. Y., Song, B. J., Kim, E. R., Bae, Y. S. & Moon, Y. H. (2000), 'Truncated form of importin alpha identified in breast cancer cells inhibits nuclear import of p53', *J. Biol. Chem.* **275**, 23139–23145.

Kim, J. K. (2015), 'Protein sequestration versus hill-type repression in circadian clock models', *IET Sys. Biol.* **10**, 125–135.

Kobayashi, T. & Kageyama, R. (2011), 'Hes1 oscillations contribute to heterogeneous differentiation responses in embryonic stem cells', *Genes* **2**, 219–228.

Kobayashi, T., Mizuno, H., Imayoshi, I., Furusawa, C., Shirahige, K. & Kageyama, R. (2009), 'The cyclic gene Hes1 contributes to the diverse differentiation responses of embryonic stem cells', *Genes and Development* **23**, 1870–1875.

Kuhlman, T. E. & Cox, E. C. (2012), 'Gene location and dna density determine transcription factor distributions in escherichia coli', *Mol Sys Biol* **8**(610).

Kyriakis, J. M. (2007), 'The integration of signalling by multiprotein complexes containing Raf kinases', *Biochimica et Biophysica Acta* **1773**, 1238–1247.

Lahav, G., Rosenfeld, N., Sigal, A., Geva-Zatorsky, N., Levine, A. J., Elowitz, M. B. & Alon, U. (2004), 'Dynamics of the p53-Mdm2 feedback loop in individual cells', *Nature* **36**, 147–150.

Lander, E., Page, D. & Lifton, R. (2004), *Annual reviews of genomics and human genetics*, Annual Reviews.

Lee, R. E. C., Walker, S. R., Savery, K., Frank, D. A. & Gaudet, S. (2014), 'Fold change of nuclear NF$-\kappa$B determines TNF-induced transcription in single cells', *Molecular Cell* **10**.

Lewis, J. (2003), 'Autoinhibition with transcriptional delay: A simple mechanism for the zebrafish somitogenesis oscillator.', *Curr. Bio.* **13**, 1398–1408.

Lipniacki, T., Paszek, P., Brasier, A. R., Luxon, B. & Kimmel, M. (2004), 'Mathematical model of the NFκB regulatory module', *Journal of Theoretical Biology* **228**, 195–215.

Lohmann, J., D'Huys, O., Haynes, N. D., Scholl, E. & Gauthier, D. J. (2017), 'Transient dynamics and their control in time-delay autonomous boolean ring networks', *Phys. Rev. E* **95**(2).

Mackey, M. C. & Glass, L. (1977), 'Oscillation and chaos in physiological control systems', *Science* **197**, 287–289.

Mackey, M. C., Santillan, M., Tyran-Kaminska, M. & Zeron, E. S. (2015), 'The utility of simple mathematical models in understanding gene regulatory dynamics.', *In Silico Biol.* **12**(1-2), 23–53.

Madian, A. G., Wheeler, H. E., Jones, R. B. & Dolan, M. E. (2012), 'Relating human genetic variation to variation in drug responses', *Special Issue: Human Genetics* **28**, 487–495.

Mahaffy, J. M. (1984), 'Models of genetic control by repression with time delays and spatial effects', *J. Mat. Biol.* **20**, 39–57.

Mahaffy, J. M. (1988), 'Genetic control models with diffusion and delays.', *Math. Biosci.* **90**, 519–533.

Mallat, S. A. (1998), *A wavelet tour of signal processing*, Academic Press.

Marquez-Lago, T. T. & Stelling, J. (2010), 'Counter-intuitive stochastic behaviour of simple gene circuits with negative feedback', *Biophys. J.* **98**, 1742–1750.

Mather, W., Bennett, M. R., Hasty, J. & Tsimring, L. S. (2009), 'Delay-induced degrade-and-fire oscillations in small genetic circuits', *Phys. Rev. Lett.* **102**.

Maxwell, P. J., Gallagher, R., Seaton, A., Wilson, C., Scullin, P., Pettigrew, J., Stratford, I. J., Williams, K. J., Johnston, P. G. & Waugh, D. J. J. (2007), 'Hif-1 and nf-κb-mediated upregulation of cxcr1 and cxcr2 expression promotes cell survival in hypoxic prostate cancer cells', *Oncogene* **26**, 7333–7345.

Maynard, M. A., Qi, H., Chung, J., Lee, E. H. L., Kondo, Y., Hara, S., Conway, R. C., Conaway, J. W. & Ohh, M. (2003), 'Multiple splice variants of the human hif-3α locus are targets of the von hippel-lindau e3 ubiquitin ligase complex', *J. Biol. Chem* **278**, 11032–11040.

Meister, A., Du, C., Li, Y. H. & Wong, W. H. (2014), 'Modeling stochastic noise in gene regulatory systems.', *Quant. Biol.* **2**(1), 1–29.

Mengel, B., Hunziker, A., Pedersen, L., Trusina, A., Jensen, M. H. & Krishna, S. (2007), 'Modeling oscillatory control in nf-κb, p53 and wnt signaling', *Current opinions in genetics and development* **20**, 1–9.

Meyers, J., Craig, J. & Odde, D. J. (2006), 'Potential for control of signaling pathways via cell size and shape', *Current Biology* **16**, 1686–1693.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D. & Alon, U. (2002), 'Network motifs: Simple building blocks of complex networks', *Science* **298**, 824–827.

Mitari, N., Jensen, M. H. & Semsey, S. (2015), 'Coupled positive and negative feedbacks produce diverse gene expression patterns in colonies', *American Society for Microbiology* **6**, 1–10.

Mohan, C. (2009), *Signal Transduction: A short overview of its role in health and disease*, Merck.

Momiji, H. & Monk, N. A. M. (2008), 'Dissecting the dynamics of the hes1 genetic oscillator', *J. Theor. Biol.* **254**, 784–798.

Monk, N. A. M. (2003), 'Oscillatory expression of Hes1, p53 and NFκB driven expression by transcriptional time delays', *Current Biology* **13**, 1409–1413.

Morriss-Kay, G. M. & Sakolova, N. (1996), 'Embryonic development and pattern formation', *FASEB J.* **10**(9), 961–968.

Muller, S., Hofbauer, J., Endler, L., Flamm, C., Widder, S. & Schuster, P. (2006), 'A generalised model of the repressilator', *Journal of Mathematical Biology* **53**, 905–937.

Naqib, F., Quail, T., Musa, L., Vulpe, H., Nadeau, J., Lei, J. & Glass, L. (2012), 'Tunable oscillations and chaotic dynamics in systems with localised synthesis.', *Phys. Rev. E* **85**(046210).

Nelson, D. E., Ihekwaba, A. E. C., Elliott, M., Johnson, J. R., Gibney, C. A., Foreman, B. E., Nelson, G., See, V., Horton, C. A., Spiller, D. G., Edwards, S. W., McDowell, H. P., Unitt, J. F., Sullivan, E., Grimley, R., Benson, N., Broomhead, D., Kell, D. B. & White, M. R. H. (2004), 'Oscillations in NF-κB signaling control the dynamics of gene expression', *Science* **306**(5696), 704–708.

Neves, S. R., Tsokas, P., Sarkar, A., Grace, E. A., Rangamani, P., Taubenfeld, S. M., Alberini, C. M., Schaff, J. C., Blitzer, R. D., Moraru, I. I. & Iyengar, R. (2008), 'Cell shape and negative links in regulatory motifs together control spatial information flow in signaling networks', *Cell* **133**, 666–680.

Norvell, A., Debec, A., Finch, D., Gibson, L. & Thoma, B. (2005), 'Squid is required for efficient posterior localization of oskar mrna during drosophila oogenesis', *Dev. Genes Evol.* **215**, 340–349.

O'Brien, E. L., Itallie, E. V. & Bennet, M. R. (2012), 'Modeling synthetic gene oscillators', *Mathematical Biosciences* **236**, 1–15.

O'Dea, E. & Hoffmann, A. (2010), 'The regulatory logic of the NFκB signaling system', *Cold Spring Harbor: Perspectives in Biology* **10**.

Oliveira, S. M. D., Chandrasseelan, J. G., Hakkinen, A., Goncalves, N. S. M., Yli-Harja, O., Startceva, S. & Ribeiro, A. S. (2015), 'Single-cell kinetics of a repressilator when implemented in a single-copy plasmid', *Molecular BioSystems* **11**, 1939–1945.

Paszek, P., Jackson, D. A. & White, M. R. H. (2010), 'Oscillatory control of signalling molecules', *Current opinions in genetics and developments* **20**, 670–676.

Potapov, I., Zhurov, B. & Volkov, E. (2015), 'Multi-stable dynamics of the non-adiabatic repressilator', *J. R. Soc. Interface* **12**(104).

Purcell, O., Savery, N. J. & Grierson, C. S. di Bernardo, M. (2010), 'A comparative analysis of synthetic genetic oscillators.', *J. R. Soc. Interface* **7**, 1503–1524.

Purvis, J. E., Karhohs, K. W., Mock, C., Batchelor, E., Loewer, A. & Lahav, G. (2012), 'p53 dynamics control cell fate', *Science* **336**, 1440–1444.

Purvis, J. E. & Lahav, G. (2013), 'Encoding and decoding cellular information through signalling dynamics', *Cell* **152**, 945–956.

Raser, J. M. & O'Shea, E. K. (2005), 'Noise in gene expression: Origins, consequences and control', *Science* **309**.

Renner, F. & Schmitz, M. L. (2009), 'Autoregulatory feedback loops terminating the nfκb response', *Trends Biochem. Sci* **34**, 128–135.

Ribeiro, A., Zhu, R. & Kauffman, S. A. (2006), 'A general modeling strategy for gene regulatory networks with stochastic dynamics', *J. Comput. Biol.* **13**(9).

Rius, J., Guma, M., Schachtrup, C., Akassoglou, K., Zinkernagel, A. S., Nizet, V., Johnson, R. S., Haddad, G. G. & Karin, M. (2008), 'NFκB links innate immunity to the hypoxic response through transcriptional regulation of HIF1α', *Nature* **453**, 807–812.

Roosen-Runge, F., Henning, M., Zhang, F., Jacobs, R. M. J., Sztucki, M., Schober, H., Seydel, T. & Schreiber, F. (2011), 'Protein self-diffusion in crowded solutions', *PNAS* **108**, 11815–11820.

Shen-Orr, S. S., Milo, R., Mangan, S. & Alon, U. (2002), 'Network motifs in the transcriptional regulation network of escherichia coli', *Nature Genetics* **31**, 64–68.

Shimojo, H., Ohtsuka, T. & Kageyama, R. (2008), 'Oscillations in notch signalling regulate maintenance of neural progenitors', *Neuron* **58**, 52–64.

Shymko, R. M. & Glass, L. (1974), 'Spatial switching in chemical reactions with heterogeneous catalysis', *J. Chem. Phys* **60**.

Skaug, B., Chen, J., Du, F., He, J., Ma, A. & Chen, Z. J. (2011), 'Direct, noncatalytic mechanism of ikk inhibition by a20', *Mol. Cell* **44**, 559–571.

Skorokhod, A. V., Hoppensteadt, F. C. & Salehi, H. D. (2002), *Random perturbation methods*, Springer-Verlag, New York.

Smith, H. (1987), 'Oscillations and multiple steady states in a cyclic gene model with repression', *J. Math. Biol.* **25**, 169–190.

Smolen, P., Baxter, D. A. & Byrne, J. H. (2001), 'Modeling circadian oscillations with interlocking positive and negative feedback loops.', *J. Neurosci.* **21**, 6644–6656.

Smolen, P., Baxter, D. A. & Byrne, J. H. (2002), 'A reduced model clarifies the role of feedback loops and time delays in the drosophila circadian oscillator.', *Biophys J.* **83**, 2349–2359.

Smolen, P., Baxter, D. & Byrne, J. H. (1999), 'Effects of macromolecular transport and stochastic fluctuations on the dynamics of genetic regulatory systems.', *Am. J. Physiol.* **277**, C777–C790.

Sneppen, K., Krishna, S. & Semsey, S. (2010), 'Simplified models of biological networks', *Annual Review Biophysics* **39**, 43–59.

Strelkowa, N. (2011), Stochastic analysis of nonlinear dynamics and feedback control for gene regulatory networks with applications to synthetic biology, PhD thesis, Imperial College London.

Strelkowa, N. & Barahona, M. (2010), 'Switchable genetic oscillator operating in quasi-stable mode', *J. R. Soc. Interface* **7**(48), 1071–1082.

Strelkowa, N. & Barahona, M. (2011), 'Transient dynamics around unstable periodic orbits in the generalised repressilator model', *Chaos* **21**(2).

Sturrock, M., Hellander, A., Matzavinos, A. & Chaplain, M. A. J. (2013), 'Spatial stochastic modelling of the Hes1 gene regulatory network: intrinsic noise can explain heterogeneity in embryonic stem cell differentiation', *Journal of The Royal Society* **10**(80).

Sturrock, M., Terry, A. J., Xirondimas, D. P., Thompson, A. M. & Chaplain, M. A. J. (2011*a*), 'Influence of the nuclear membrane, active transport and cell shape on the Hes1 and p53-Mdm2 pathways: Insights from spatio-temporal modelling', *Bulletin of Mathematical Biology* **74**, 1531–1579.

Sturrock, M., Terry, A. J., Xirondimas, D. P., Thompson, A. M. & Chaplain, M. A. J. (2011*b*), 'Spatio-temporal modelling of the Hes1 and p53-Mdm2 intracellular signalling pathways', *Journal of Theoretical Biology* **273**, 15–31.

Suel, G. M., Garcia-Ojalvo, J., Liberman, V. & Elowitz, M. B. (2006), 'An excitable

gene regulatory circuit induces transient cellular differentiation', *Nature* **440**, 545–550.

Takahashi, K., Tanase-Nicola, S. & ten Wolde, P. R. (2010), 'Spatio-temporal correlations can drastically change the response of a mapk pathway', *Proceedings of the National Academy of Sciences* **107**, 2473–2478.

Tay, S., Hughey, J. J., Lee, T. K., Lipniacki, T., Quake, S. R. & Covert, M. W. (2010), 'Single-cell nf-κb dynamics reveal digital activation and analogue information processing', *Nature* **466**, 267–272.

Terry, A. J. & Chaplain, M. A. J. (2011), 'Spatio-temporal modelling of the NF-κB intracellular signalling pathway: The roles of diffusion, active transport, and cell geometry', *Journal of Theoretical Biology* **290**, 7–26.

Thompson, D. (1917), *On growth and form: The complete revised edition (1992)*, Dover Publics., New York.

Thuong, N. T., Dunstan, S. J., Chau, T. T. H., Thorsson, V., Simmons, C. P., Quyen, N. T. H., Thwaites, G. E., Lan, N. T. N., Hibberd, M., Teo, Y. Y., Seielstad, M., Aderem, A., Farrar, J. J. & Hawn, T. R. (2008), 'Identification of tuberculosis susceptibility genes with human macrophage gene expression profiles', *PLoS Pathogens* **4**, 1–13.

Tian, T. & Burrage, B. (2006), 'Stochastic models for regulatory networks of the genetic toggle switch', *Proc. Nat. Acad. Sci.* **103**, 8372–8377.

Tian, T., Burrage, B., Burrage, P. M. & Carletti, M. (2007), 'Stochastic delay differential equations for genetic regulatory networks', *J. Comp. Appl. Math.* **205**, 696–707.

Tian, X., Zhang, X., Liu, F. & Wang, W. (2009), 'Interlinking positive and negative feedback loops creates a tunable motif in gene regulatory networks', *Physical Review E* **80**, 1–8.

Tiana, G., Jensen, M. H. & Sneppen, K. (2002), 'Time delay as a key to apoptosis induction in the p53 network', *The European Physics Journal B* **29**, 135–140.

Torrence, C. & Compo, G. P. (1998), 'A practical guide to wavelet analysis', *Bull. Am. Meteorol. Soc.* **79**, 61–78.

Turcotte, M., Garcia-Ojalvo, J. & Suel, G. M. (2008), 'A genetic timer through noise-induced stabilization of an unstable state', *Proc. Natl Acad. Sci.* **105**, 15732–15737.

Turing, A. M. (1952), 'The chemical basis of morphogenesis', *Philos. Trans. Roy. Soc. London* **237**(641), 37–72.

Turner, D. A., Paszek, P., Woodcock, D. J., Nelson, D. E., Horton, C. A., Wang, Y., Spiller, D. G., Rand, D. A., White, M. R. H. & Harper, C. V. (2010), 'Physiological levels of TNFα stimulation induce stochastic dynamics of NF$-\kappa$B response in single living cells', *Journal of Cell Science* **123**, 2834–2843.

van Uden, P., Kenneth, N. S. & Rocha, S. (2008), 'Regulation of hypoxia-inducible factor-1α by nf-κb', *J. Biochem* **412**, 477–484.

van Zon, J. S., Morelli, M. J., Tanase-Nicola, S. & ten Wolde, P. R. (2006), 'Diffusion of transcription factors can drastically enhance the noise in gene expression', *Biophys. J.* **91**, 4350–4367.

Veening, J., Smits, W. K. & Kuipers, O. P. (2008), 'Bistability, epigenetics, and bet-hedging in bacteria', *Annual Review Microbiology* **62**, 193–210.

Wang, J., Zhang, J., Yuan, Z. & Zhou, T. (2007), 'Noise-induced switches in networks systems of the genetic toggle switch', *BMC Sys. Biol.* **1**(50).

Wang, N., Butler, J. P. & Ingber, D. E. (1993), 'Mechanotransduction across the cell surface and through the cytoskeleton', *Science* **260**, 1124–1127.

Wang, Y., Ha, U., Zeng, L. & Jin, S. (2003), 'Regulation of membrane permeability by a two component regulatory system in pseudomonas aeruginosa', *Antimicrobial, Agents and Chemotherapy* pp. 95–101.

Wang, Y., Paszek, P., Horton, C. A., Kell, D. B., White, M. R. H., Broomhead, D. S. & Muldoon, M. R. (2011), 'Interactions among oscillatory pathways in nf-κb signaling', *BMC Systems Biology* **5**, 1–11.

Warren, P. B. & ten Wolde, P. R. (2005), 'Chemical models of genetic toggle switches', *J. Phys. Chem. B.* **109**(14).

Wawra, C., Kuhl, M. & Kestler, H. A. (2007), 'Extended analysis of the wnt/β-catenin pathway: Robustness and oscillatory behaviour', *FEBS Letters* **581**, 4043–4048.

Wirtz, H. R. & Dobbs, L. G. (1990), 'Calcium mobilisation and exocytosis after one mechanical stretch of lung epithelial cells', *Science* **250**, 1266–1269.

Xu, Y., Zhu, Y., Shen, J. & Su, J. (2014), 'Switch dynamics for stochastic model of genetic toggle switch', *Physica : Statistical mechanics and its applications* **416**, 461–466.

Yeger-Lotem, E., Sattath, S., Kashtan, N., Itzkovitz, S., Milo, R., Pinter, R. Y., Alon, U. & Margalit, H. (2004), 'Network motifs in integrated cellular networks of transcription-regulation and protein-protein interaction', *Proceedings of the National Academy of Sciences* **101**, 5934–5939.

Yildirim, N. & Mackey, M. C. (2003), 'Feedback regulation in the lactose operon: A mathematical modeling study and comparison with experimental data.', *Biophys. J.* **84**, 2841–2851.

Yuh, C. H., Bolouri, H. & Davidson, E. H. (2001), 'Cis-regulatory logic in the endo16 gene: switching from a specification to a differentiation mode of control', *Development* **128**, 617–629.

Zeng, T. & Chen, L. (2012), 'Tracing dynamic biological processes during phase transitions', *BMC Systems Biology* **6**, 1–14.